ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

ON THE THEORY OF LINEAR INTEGRAL EQUATIONS. II[1]

BY EINAR HILLE AND J. D. TAMARKIN

(Received March 21, 1934)

I. On completely continuous integral transformations

1. Let L_p be the space of functions $f(s)$ measurable over $(-\infty, \infty)$ and such that the integral $\int_{-\infty}^{\infty} |f(s)|^p\, ds < \infty$. We set

$$\|f\|_p = \left(\int_{-\infty}^{\infty} |f(s)|^p\, ds\right)^{1/p}, \tag{1}$$

taking this for the definition of a metric of the space L_p. We shall write

$$f_h(s) = 1/h \int_s^{s+h} f(t)\, dt, \tag{2}$$

$$f^N(s) = \begin{cases} f(s) & \text{if} \quad |s| \leqq N, \\ 0 & \text{elsewhere.} \end{cases} \tag{3}$$

Let $\mathfrak{E}$ be a sub-set of the space L_p. Necessary and sufficient conditions that $\mathfrak{E}$ shall be compact, in the case when $p > 1$ and the functions of $\mathfrak{E}$ vanish outside a fixed interval, were found by Kolmogoroff [7]. They reduce to

(i) $\|f\|_p \leqq A$, where A is a constant depending only on $\mathfrak{E}$,

(ii) $\|f - f_h\|_p \to 0$ as $h \to 0$, uniformly over $\mathfrak{E}$.

It was shown by Tamarkin [10] that if the restriction of vanishing outside a fixed interval be dropped, a further condition should be added,

(iii) $\|f - f^N\|_p \to 0$ as $N \to \infty$, uniformly over $\mathfrak{E}$,

constituting together with (i) and (ii) a complete set of necessary and sufficient conditions for the compactness of $\mathfrak{E}$ in L_p. As shown by Tulajkov [11] this result holds even when $p = 1$. Finally it was shown by M. Riesz [8] that Kolmogoroff's condition (ii) may be replaced by

(ii′) $\|f(s) - f(s + h)\|_p \to 0$ as $h \to 0$, uniformly over $\mathfrak{E}$,

which is conveniently characterized by saying that the functions of $\mathfrak{E}$ are equally continuous in L_p. Condition (ii) follows from (ii′) by an immediate

[1] This note should be considered as a supplement to our previous papers [4, 5]. The numbers in brackets refer to the bibliography at the end of this note. The result of Part II was stated without proof in our paper [5, p. 6]. The result of Part I was obtained but not published, in October 1931.

application of Hölder's inequality. In the present note we shall use the set of conditions (i), (ii′) and (iii).

2. Let $L_{\alpha,\beta}$ designate the class of kernels defined and measurable over $-\infty < s, t < \infty$, and such that

$$\int_{-\infty}^{\infty} ds \left(\int_{-\infty}^{\infty} | K(s,t) |^{\beta} \, dt \right)^{\alpha/\beta} < \infty .$$

We shall be particularly interested in the class $L_{p,p'}$, where $p > 1$ and $p' = p/(p-1)$. In this case we set for abbreviation

$$l(s;K) \equiv l(s) = \left(\int_{-\infty}^{\infty} | K(s,t) |^{p'} \, dt \right)^{1/p'}, \tag{4}$$

$$|| K || = || l(s;K) ||_p = \left(\int_{-\infty}^{\infty} l(s;K)^p \, ds \right)^{1/p}, \tag{5}$$

and note the inequality

$$|| K_1 + K_2 || \leqq || K_1 || + || K_2 || \tag{6}$$

which is readily proved by a repeated application of Minkowski's inequality,

$$\begin{aligned} || K_1 + K_2 || &= || l(s;K_1 + K_2) ||_p \leqq || l(s;K_1) + l(s;K_2) ||_p \\ &\leqq || l(s;K_1) ||_p + || l(s;K_2) ||_p = || K_1 || + || K_2 ||. \end{aligned}$$

The result of the present note is embodied in the following

THEOREM. *The integral transformation represented by*

$$y(s) = K \cdot x(s) = \int_{-\infty}^{\infty} K(s,t) \, x(t) \, dt \tag{7}$$

is a completely continuous[2] *transformation on L_p to L_p, $p > 1$.*

We first observe that if $x = x(s) \subset L_p$, then $y = K \cdot x$ also $\subset L_p$. This follows at once by a repeated application of Hölder's inequality,

$$| y(s) | \leqq || x ||_p \, l(s;K), \tag{8}$$

$$|| y ||_p \leqq || K || \, || x ||_p . \tag{9}$$

To prove the theorem we have only to show that if $\mathfrak{F}$ is any bounded sub-set of L_p then the K-transform $\mathfrak{E} = K \cdot (\mathfrak{F})$ of $\mathfrak{F}$ is compact. To do this we apply the set of conditions (i), (ii′), (iii) of §1. If we assume that

$$|| x ||_p \leqq A, \qquad x \subset \mathfrak{F}, \tag{10}$$

---

[2] We refer to the monograph by Banach [1] for the terminology and bibliography of the theory of linear operations used in the present note.

then, by (9),

$$\| y \|_p \leqq A \| K \|, \qquad y \subset \mathfrak{E},$$

which shows that (i) is satisfied.

To prove (iii) we note that, by (8) and (10),

$$\| y - y^N \|_p^p = \int_{-\infty}^{-N} | y(s) |^p ds + \int_N^{\infty} | y(s) |^p ds$$

$$\leqq A^p \left[\int_{-\infty}^{-N} [l(s; K)]^p ds + \int_N^{\infty} [l(s; K)]^p ds \right] \to 0 \text{ as } N \to \infty,$$

uniformly over $\mathfrak{E}$.

Condition (ii′) requires a slightly longer discussion. We have, on setting

$$K_h(s, t) = K(s + h, t), \tag{11}$$

$$y(s + h) - y(s) = (K_h - K) \cdot x(s),$$

$$\| y(s + h) - y(s) \|_p \leqq \| x \|_p \| K_h - K \| \leqq A \| K_h - K \|.$$

Hence (ii′) will be verified provided we can show that

$$\| K_h - K \| \to 0 \quad \text{as} \quad h \to 0. \tag{12}$$

This can be achieved in three steps. First we show that if

$$K^N(s, t) = \begin{cases} K(s, t), & \text{if} \quad | s | \leqq N, | t | \leqq N, \\ 0 & \text{elsewhere,} \end{cases} \tag{13}$$

then

$$\| K - K^N \| \to 0 \quad \text{as} \quad N \to \infty. \tag{14}$$

Indeed, we have

$$[l(s; K - K^N)]^{p'} = \begin{cases} \displaystyle\int_{-\infty}^{-N} | K(s, t) |^{p'} dt + \int_N^{\infty} |K(s, t) |^{p'} dt, & \text{if} \quad | s | \leqq N, \\ \displaystyle\int_{-\infty}^{\infty} | K(s, t) |^{p'} dt, & \text{if} \quad | s | > N. \end{cases}$$

It is now plain that, for almost all s,

$$[l(s; K - K^N)]^p \to 0 \text{ as } N \to \infty; \qquad [l(s; K - K^N)]^p \leqq [l(s; K)]^p.$$

Hence, by Lebesgue's theorem,

$$\| K - K^N \|^p = \int_{-\infty}^{\infty} [l(s; K - K^N)]^p ds \to 0 \qquad \text{as} \qquad N \to \infty.$$

As the second step we prove that if $K(s, t)$ vanishes outside a fixed square $-N \leqq s, t \leqq N$, then

$$||K - K(P)|| \to 0 \quad \text{as} \quad P \to \infty, \tag{15}$$

where

$$K(P) \equiv K(s, t; P) = \begin{cases} K(s, t), & \text{if} \quad |K(s, t)| \leqq P, \\ 0 & \text{elsewhere.} \end{cases} \tag{16}$$

To do this we introduce the set $E_P \equiv \underset{s,t}{E}\{|K(s, t)| \leqq P\}$ and its complementary, $D_P = CE_P$. Let $D_P(s_0)$ be the cross-section of D_P by the line $s = s_0$. Then, for almost all s,

$$[l(s; K - K(P))]^p = \left(\int_{D_P(s)} |K(s, t)|^{p'} dt\right)^{p/p'} \to 0 \text{ as } P \to \infty, \qquad l(s; K - K(P)) \leqq l(s; K),$$

which yields (15).

As the third step we prove that if $K(s, t)$ is bounded and vanishes outside a fixed square $-N \leqq s, t \leqq N$, then

$$||K - K'(\epsilon)|| \to 0 \quad \text{as} \quad \epsilon \to 0, \tag{17}$$

where

$$K'(\epsilon) \equiv K'(s, t; \epsilon) = 1/\epsilon \int_s^{s+\epsilon} K(\sigma, t)\, d\sigma. \tag{18}$$

We put

$$\Delta(s) \equiv \Delta(s; K) = \int_{-\infty}^{\infty} |K(s, t) - K'(s, t; \epsilon)|\, dt.$$

By a familiar argument

$$\int_{-\infty}^{\infty} \Delta(s)\, ds = \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} dt \left| 1/\epsilon \int_0^{\epsilon} [K(s + \sigma, t) - K(s, t)]\, d\sigma \right|$$

$$\leqq 1/\epsilon \int_0^{\epsilon} d\sigma \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} |K(s + \sigma, t) - K(s, t)|\, ds \to 0 \quad \text{as} \quad \epsilon \to 0.$$

But for the kernels of the class under consideration,

$$[l(s; K - K'(\epsilon))]^{p'} = \int_{-\infty}^{\infty} |K(s, t) - K'(s, t; \epsilon)|^{p'} dt \leqq C\Delta(s),$$

where C is a constant (not necessarily the same in this and in the following formulas) whose precise value is of no interest, but which does not depend on ϵ. Now, if $p > 2$, $p' < 2 < p$, and

$$\| K - K(\epsilon) \|^p = \int_{-\infty}^{\infty} [l(s; K - K'(\epsilon))]^p \, ds$$

$$\leqq C \int_{-\infty}^{\infty} [l(s; K - K'(\epsilon)]^{p'} \, ds \leqq C \int_{-\infty}^{\infty} \Delta(s) \, ds \to 0 \qquad \text{as} \qquad \epsilon \to 0.$$

If, however, $p \leqq 2 \leqq p'$, since $l(s; K - K'(\epsilon))$ vanishes outside the fixed interval $|s| \leqq N + 1$, as soon as $\epsilon < 1$, we have by Hölder's inequality

$$\int_{-\infty}^{\infty} [l(s; K - K'(\epsilon))]^p \, ds$$

$$\leqq (2N + 2)^{2-p} \left(\int_{-\infty}^{\infty} [l(s; K - K'(\epsilon))]^{p'} \, ds \right)^{p/p'} \leqq C \left(\int_{-\infty}^{\infty} \Delta(s) \, ds \right)^{p/p'},$$

and again

$$\| K - K'(\epsilon) \| \to 0 \qquad \text{as} \qquad \epsilon \to 0.$$

As the last step we observe that (12) is trivial in the case where $K(s, t)$ is bounded and vanishes outside a fixed square and is continuous in s for almost all t. On combining the preceding steps with the inequality (6) we complete the proof of our theorem.

3. **Remarks.** (a) The theory of completely continuous transformations as developed by F. Riesz, Hildebrandt, and Banach, shows that the classical results of the Fredholm theory are valid for the integral equation

$$x(s) = y(s) + \lambda K \cdot x(s); \qquad x, y \subset L_p. \tag{19}$$

(b) Banach [1, pp. 98–99] obtained a result analogous to our theorem above, viz., that the transformation $y = K \cdot x$ on L_p to L_q, for the interval (0, 1), is completely continuous if

$$\int_0^1 \int_0^1 |K(s, t)|^{r'} \, ds \, dt < \infty,$$

where $p > 1$, $q > 1$, and r is the smaller of the numbers p and q'. To compare with our theorem we have to take $q = p$. Then $r = \min(p, p') \leqq 2$, $r' \geqq 2$, and $K \subset L_2$. There exist, however, kernels $K \subset L_{p, p'}$, $p > 2$ (even vanishing outside a fixed square) which do not $\subset L_2$. It suffices to take [4, p. 482] $K(s, t) = a_n \geqq 0$ if $b_{n-1} \leqq s, t \leqq b_n$, $n = 1, 2, 3, \cdots$; $b_0 = 0, b_n \uparrow b < \infty$, and $K(s, t) = 0$ elsewhere. It is readily found that, for such kernels,

$$\| K \|^p = \sum_{n=1}^{\infty} a_n^p \, \delta_n^p, \qquad \delta_n = b_n - b_{n-1},$$

while

$$\int_0^1 \int_0^1 K(s, t)^2 \, ds \, dt = \sum_{n=1}^{\infty} a_n^2 \, \delta_n^2,$$

and it remains only to select a_n and δ_n so that

$$\sum_n a_n^p \, \delta_n^p < \infty, \qquad \sum_n \delta_n = b, \qquad \sum_n a_n^2 \, \delta_n^2 = \infty$$

which is obviously always possible. It is worth noticing that the C.V. of our kernel K are $\lambda_n = (a_n \delta_n)^{-1}$ and that the series $\Sigma \lambda_n^{-p}$ converges. Whether this property holds for the general kernel $K \subset L_{p,\, p'}$ is an open question.

(c) A natural interpretation of the class $L_{p,\, p'}$ in the limiting cases $p = \infty$, $p = 1$ is given by

$$(L_{\infty,\, 1}) \qquad \int_{-\infty}^{\infty} | K(s, t) | \, dt \leqq A < \infty, \text{ for almost all } s,$$

$$(L_{1,\, \infty}) \qquad \int_{-\infty}^{\infty} \max_t | K(s, t) | \, ds < \infty,$$

where max means essential least upper bound. Our theorem fails when $p = 1$ or $p = \infty$. In case $p = \infty$ this is shown by the example of the kernel

$$K(s, t) = \begin{cases} s^{-1}, \text{ if } 0 < t \leqq s, \\ 0 \qquad \text{elsewhere,} \end{cases}$$

for which all λ with $\Re(\lambda) \geqq 1$ are C.V., with the corresponding (bounded) fundamental functions $s^{\lambda - 1}$ [4, pp. 525–527].

That our theorem fails in the case $p = 1$ is shown by the existence of bounded measurable kernels (vanishing outside of the square $0 \leqq s, t \leqq 1$) such that the transformation $K \cdot x$ is not completely continuous in L. The following example of such a kernel was communicated to us by Professor J. v. Neumann. We introduce an orthonormal set $\{\varphi_\nu(s)\}$ defined over (0, 1) as follows

$$\varphi_{2n-i}(s) = \begin{cases} 2^{\frac{1}{2}} \sin 2n\pi s, & 0 < s < \frac{1}{2}, \\ (-1)^{i+1} 2^n, & 1 - 2^{-2n} < s < 1 - 2^{-2n-1}, \\ 0 \text{ elsewhere.} \end{cases} \qquad \begin{pmatrix} n = 1, 2, \cdots, \\ i = 0, 1. \end{pmatrix}$$

We put

$$K(s, t) = \sum_{n=1}^{\infty} 2^{-n} [\varphi_{2n-1}(s) \, \varphi_{2n-1}(t) - \varphi_{2n}(s) \, \varphi_{2n}(t)].$$

It is readily seen that

$$| K(s, t) | \leqq 2^{\frac{1}{2}},$$

so the kernel is bounded. It is plain that

$$K \cdot \varphi_{2n-i}(s) = (-1)^{i+1} 2^{-n} \varphi_{2n-i}(s), \qquad i = 0, 1,$$

whence, upon setting

$$\psi_n(s) = 2^n [\varphi_{2n-1}(s) - \varphi_{2n}(s)],$$

we deduce

$$K \cdot \psi_n(s) = \varphi_{2n-1}(s) + \varphi_{2n}(s).$$

On the other hand

$$|| \psi_n ||_1 = 2^n \int_0^1 | \varphi_{2n-1}(s) - \varphi_{2n}(s) | \, ds = 1,$$

while, for $n \neq m$

$$\begin{aligned} || K \cdot \psi_n - K \cdot \psi_m ||_1 &= \int_0^1 | [\varphi_{2n-1}(s) + \varphi_{2n}(s)] - [\varphi_{2m-1}(s) + \varphi_{2m}(s)] | \, ds \\ &\geqq 2^{\frac{1}{2}} \int_0^{\frac{1}{2}} | \sin 2n\pi s - \sin 2m\pi s | \, ds \\ &= 4 \cdot 2^{\frac{1}{2}} \int_0^{\frac{1}{2}} | \sin (n - m) \pi s | \, | \cos (n + m) \pi s | \, ds \\ &\geqq 4 \cdot 2^{\frac{1}{2}} \int_0^{\frac{1}{2}} \sin^2 (n - m) \pi s \cos^2 (n + m) \pi s \, ds \\ &= 2^{-\frac{1}{2}}. \end{aligned}$$

Thus the transformation $K \cdot x$ carries the sequence $\{\psi_n(s)\}$, which is bounded in L, into the sequence $\{K \cdot \psi_n\}$ which does not contain any convergent subsequence. This transformation is consequently not completely continuous in L.

Suppose, however, that condition $(L_{1,\infty})$ is satisfied, and let

$$\int_{-\infty}^{\infty} \max_t | K(s, t) | \, ds = A.$$

Further let $E(t, N)$ be the set of values of s for which $| K(s, t) | > N$. Then for all values of t, except possibly for a set D of measure zero, we have meas $E(t, N) \leqq A/N$. Thus for all $t \subset CD$,

$$\int_{-\infty}^{\infty} | K(s, t) - K^N(s, t) | \, ds = \int_{E(t, N)} | K(s, t) | \, ds$$

tends to zero as $N \to \infty$, uniformly over CD, since, for such values of t, $K(s, t)$ is dominated by the fixed integrable function $\max_t | K(s, t) |$ while meas $E(t, N) \to 0$ uniformly over CD. This shows that condition $(L+)$ of our paper [4, p. 509] is satisfied. Hence the classical results of the Fredholm theory still hold in the

case $(L_{1,\infty})$ although the transformation $y = K \cdot x$ in general is not completely continuous in L.

(d) Professor Hildebrandt was kind enough to communicate to us another method of proving the theorem of §3, of which he has been in possession since 1930 and which is based upon the possibility of approximating the given kernel $K(s, t) \subset L_{p,p'}$ by kernels of finite rank $K_n(s, t) = \sum_1^n \psi_i(s)\, \varphi_i(t)$ in such a way that $\| K - K_n \| \to 0$. Since the transformation $y = K_n \cdot x$ is obviously completely continuous in L_p, it follows at once that the transformation $y = K \cdot x$ is also completely continuous. Assume for the sake of simplicity that $K(s, t)$ vanishes outside the square $0 \leqq s, t \leqq 1$. Then we may take for $\{\varphi_\nu(s)\}$ the orthonormal set of Haar, with $\psi_\nu(s) = K \cdot \varphi_\nu(s)$. Since $1 < p' < \infty$, we have [see 9], for almost all s,

$$\int_0^1 | K(s, t) - K_n(s, t) |^{p'}\, dt \to 0 \qquad \text{as} \qquad n \to \infty,$$

$$\int_0^1 | K_n(s, t) |^{p'}\, dt \leqq \int_0^1 | K(s, t) |^{p'}\, dt.$$

Hence, for almost all s,

$$l(s; K - K_n) \to 0 \qquad \text{as} \qquad n \to \infty,$$

while

$$l(s; K - K_n)^p \leqq [l(s; K) + l(s; K_n)]^p \leqq [2l(s; K)]^p.$$

Consequently

$$\| K - K_n \|^p = \int_0^1 l(s; K - K_n)^p\, ds \to 0 \qquad \text{as} \qquad n \to \infty.$$

The extension to the infinite range does not offer any essential difficulty.

(e) Finally we call attention to the analogy between our result and those of Bóbr [2] and Cohen [3] in the case of system of infinitely many equations with infinitely many unknowns.

II. On the exponent of convergence of characteristic values of certain kernels

1. Let $K(s, t)$ be bounded and measurable over $a \leqq s, t \leqq b$ (a and b finite). The resolvant $R(s, t; \lambda; K) \equiv R(s, t) \equiv R$ of K is determined uniquely (almost everywhere) by the equations

$$K(s, t) + R(s, t) = \lambda K \cdot R(s, t) = \lambda R \cdot K(s, t).^3 \tag{1}$$

---

[3] We use the convenient symbolic notation $A \cdot B = A \cdot B(s, t) = \int_a^b A(s, u)\, B(u, t)\, du$,

$A \cdot x \equiv A \cdot x(s) = \int_a^b A(s, t)\, x(t)\, dt$, $\| A \|^2 = \int_a^b \int_a^b | A(s, t) |^2\, ds\, dt$.

This is a meromorphic function of λ represented by the ratio of two entire functions

$$R(s, t; \lambda; K) = \frac{D(s, t; \lambda; K)}{D(\lambda; K)}, \tag{2}$$

where $D(\lambda; K)$ is the Fredholm determinant of K. The characteristic values $\lambda_1, \lambda_2, \lambda_3, \cdots$ (C.V.) of K are the zeros of $D(\lambda; K)$. If each of them is repeated according to its multiplicity as a zero of $D(\lambda; K)$ then we have the classical inequality of Schur

$$\sum_{\nu=1}^{\infty} |\lambda_\nu|^{-2} \leqq ||K||^2. \tag{3}$$

It is well known that these results still hold if the assumption that K is bounded is replaced by the assumption $K \subset L_2$, which means to say that the right-hand member of (3) is finite. In this case, however, the Fredholm determinant may not exist and should be replaced either by the "modified Fredholm determinant" $D^*(\lambda; K)$ of Hilbert-Carleman, or by a certain infinite determinant $A^*(\lambda; K)$ with the corresponding modification of the formula (2) for the resolvant.[4] Formula (3) still holds provided the λ_ν are defined as the zeros of $D^*(\lambda; K)$ or of $A^*(\lambda; K)$ respectively, each repeated according to its multiplicity. This definition of the "multiplicity" of a C.V. of K has the apparent disadvantage of being dependent on the particular representation of the resolvant R. However, according to the well known Lalesco theory of the principal part of the resolvant corresponding to a C.V. λ_0, the multiplicity of λ_0 as defined above coincides with the total number of pairs of the "principal functions" of K associated with λ_0.[5] This furnishes a more satisfactory definition of the multiplicity of a C.V. of K in as much as it is entirely determined by R (hence by K) itself, and does not depend on any special representation of R. We shall agree therefore to *define the multiplicity of a C.V. λ_0 of K as the number of pairs of the principal functions of K, associated with λ_0, provided that the Lalesco theory can be applied to the kernel K in question.* If this is the case, we shall use the notation $\{\lambda_\nu\}$ for the sequence of the C.V. of K, each being repeated according to its multiplicity. The exponent of convergence of the sequence $\{\lambda_\nu\}$ will be denoted by $\rho(K)$.

2. There are numerous cases where the resolvant of K, as defined by (1), exists though K is neither bounded nor $\subset L_2$. One such case was investigated by Hobson [6],[6] under the main assumption that there exists an integer m such

[4] See, e.g. [5] where other bibliographical references are found.

[5] Although only the case of K bounded seems to have been treated in detail in the literature, there is no difficulty in extending the Lalesco theory to the case where $K \subset L_2$.

[6] It should be observed that Hobson's assumptions are not quite clearly stated. The argument of Hobson is easily made precise, however.

that in the sequence of the iterated kernels

$$K, \qquad K^2 = K \cdot K, \qquad K^3 = K \cdot K \cdot K, \cdots$$

all kernels K^m, K^{m+1}, $\cdots$ are bounded. From Hobson's discussion it follows that the resolvant R exists and that the Lalesco theory can be applied in his case. Furthermore if μ_0 is a C.V. of K^m of multiplicity r_0 then r_0 is equal to the sum of multiplicities of those among the m^{th} roots of μ_0 which are C.V. of K. Since, in view of (3)

$$\rho(K^m) \leqq 2,$$

it follows that

$$\rho(K) \leqq 2m,$$

and that moreover,

$$\sum_{\nu=1}^{\infty} |\lambda_\nu|^{-2m} < \infty. \tag{4}$$

Conditions of Hobson are satisfied in the particularly interesting case of kernels of class

$$|K(s,t)| \leqq M|s-t|^{-\alpha}, \qquad 0 \leqq \alpha < 1, \tag{H_α}$$

where M is a positive constant. A simple computation shows that

$$|K^n(s,t)| \leqq M_n|s-t|^{-\alpha_n}, \qquad M_n|\log|s-t||, \qquad M_n, \tag{5}$$

according as $\alpha_n = n\alpha - n + 1$ is > 0, $= 0$, or < 0. Here M_n is a constant analogous to M. We may therefore take $m = [(1-\alpha)^{-1}] + 1$. This leads to the corresponding estimate $\rho(K) \leqq 2[(1-\alpha)^{-1}] + 2$. It turns out, however, that this estimate may be considerably improved by a slight modification of the preceding method.

3. Instead of Hobson's assumptions let us now assume that, for a certain value of m, K^m, K^{m+1}, $\cdots \subset L_2$. It is not difficult to formulate a set of sufficient additional conditions on K under which Hobson's argument will retain its validity. Taken in their generality, these conditions are of a fairly complicated nature and do not present considerable interest. That is why we prefer to turn instead to the most interesting special case of kernals of class H_α. A simple but tedious scrutiny of Hobson's work reveals that his method, and hence the Lalesco theory can be applied in the present case. On the other hand, in view of (5), the kernel $K^n \subset L_2$ as soon as $n\alpha - n + 1 < \frac{1}{2}$, $n > \frac{1}{2}(1-\alpha)^{-1}$. We thus are led to the following theorem

THEOREM. *If the kernel $K \subset H_\alpha$, $0 \leqq \alpha < 1$, then the exponent of convergence $\rho(K)$ does not exceed $2[\frac{1}{2}(1-\alpha)^{-1}] + 2$.*

REMARK. Take $K(s, t) = K(s - t) = F(s - t)$ where $F(\theta)$ is periodic of period 2π, and for $0 < \theta < 2\pi$ is defined by

$$F(\theta) = \sum_{n=2}^{\infty} n^{\alpha-1} e^{ni\theta}.$$

It is known (Cf. [5, pp. 10–11]) that such a kernel $\subset H_\alpha$, while $\rho(K) = (1 - \alpha)^{-1}$. This can be made as near as we please to the estimate of the theorem if $\frac{1}{2}(1 - \alpha)^{-1}$ tends to any integer $k \geqq 1$ remaining less than k.

BIBLIOGRAPHY

1. S. BANACH, *Théorie des opérations linéaires*. Warsaw, 1932.
2. S. BÓBR, *Eine Verallgemeinerung des v. Kochschen Satzes über die absolute Konvergenz der unendlichen Determinanten*. Math. Zeitschr., 10 (1921), 1–11.
3. L. W. COHEN, *A note on a system of equations with infinitely many unknowns*. Bull. Amer. Math. Soc., 36 (1930), 563–572.
4. E. HILLE AND J. D. TAMARKIN, *On the theory of linear integral equations*. I. Ann. of Math., vol. 31 (1930), 479–528.
5. E. HILLE AND J. D. TAMARKIN, *On the distribution of the characteristic values of linear integral equations*. Acta Math., vol. 57 (1931), 1–76.
6. E. W. HOBSON, *On the linear integral equation*. Proc. Lond. Math. Soc., (2) 13 (1914), 307–340.
7. A. KOLMOGOROFF, *Über Kompaktheit der Funktionenmengen bei der Konvergenz im Mittel*. Göttinger Nachrichten, 1931, 60–63.
8. M. RIESZ, *Sur les ensembles compacts de fonctions sommables*. Acta Szeged, Vol. 6 (1933), 136–142.
9. J. SCHAUDER, *Eine Eigenschaft des Haarschen Orthogonalsystem*. Math. Zeitschr., Vol. 28 (1928), 317–320.
10. J. D. TAMARKIN, *On the compactness of the space L_p*, Bull. Amer. Math. Soc., 38 (1932), 79–84.
11. A. TULAJKOV, *Zur Kompaktheit im Raum L_p*, für $p = 1$, Göttinger Nachrichten, 1933, 169–170.

NEW HAVEN, CONN. AND PROVIDENCE, R. I.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

ON THE ANALYTIC NATURE OF SURFACES OF LEAST AREA

BY E. J. MCSHANE

(Received December 4, 1933)

Introduction

There is an intimate connection between the problem of finding a minimal surface bounded by a given non-self-intersecting curve and the problem of finding a surface bounded by that curve and having least possible area. The very name "minimal surface" arises from the fact that the differential equations of a minimal surface are the Euler-Lagrange equations of the problem of least area, expressing the vanishing of the first variation of the area. Moreover, if the given curve bounds a surface of finite area (which is the only case for which the problem of least area has a meaning) the two problems have common solutions.[1]

Nevertheless, it is well known that a minimal surface need not have least area for its boundary, and also that a surface of least area need not be minimal. The following problem therefore suggests itself: Given a surface S, defined by equations

$$x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v) \qquad u^2 + v^2 \leqq 1,$$

and given a region B in the circle $u^2 + v^2 \leqq 1$. What conditions on x, y, z enable us to state that the portion of S corresponding to B is minimal?

The most general result at present in the literature is as follows: The portion of S corresponding to B is minimal if it projects orthogonally in a one-to-one way on some plane.[2] Radó asks if it be sufficient to assume that the equations $x = x(u, v)$, etc., carry B in a one-to-one way into a portion of S. One of the results of the present paper is to answer this question affirmatively.

Let us give the name "excrescence" to a portion of S, not itself a single point, which joins the rest of S in a single point (for example, a line-segment forming a spine projecting from S). Analytically, this means that there is an open subset B of the unit circle such that x, y, z are constant on the boundary of B without being constant on all of B. If S has excrescences it is clearly not minimal. But it would seem to be (and actually is) possible to remove all excrescences from S,

---

[1] Jesse Douglas, *Solution of the problem of Plateau*, Trans. Am. Math. Soc. **33** (1931), pp. 263–321. E. J. McShane, *Parameterizations of saddle surfaces, etc.*, Trans. Am. Math. Soc. **35** (1933) pp. 716–733. Tibor Radó: (1) *The problem of least area and the problem of Plateau*, Math. Zeitschrift, **32** (1930) pp. 763–796. (2) *An iterative process in the problem of Plateau*, Trans. Am. Math. Soc. **35** (1933) pp. 869–887.

[2] E. J. McShane, *On a certain inequality of Steiner*, Annals of Math., **33** (1932) pp. 125–138.

thus obtaining a surface $\bar{S}$ free from excrescences. We might then hope to prove that if S has least area for its boundary, the surface $\bar{S}$ obtained by removing excrescences from S is minimal. This is however false; the surface $x(u, v) = u$, $y(u, v) \equiv z(u, v) \equiv 0$, $0 \leqq u, v \leqq 1$ has no excrescences and is not minimal. But the boundary of this surface is not a Jordan curve. We may therefore try to prove that if S is bounded by a Jordan curve Γ and has least area among all surfaces bounded by Γ, then the surface $\bar{S}$ is minimal. The proof of this statement is the principal object of the present paper. The result may also be stated thus: For a given Jordan curve Γ, the most general surface of least area bounded by Γ is a minimal surface plus excrescences. It follows that a surface of least area with a Jordan boundary and without excrescences is minimal.

This does not answer Radó's question, for it requires that the boundary Γ of the whole surface S be a Jordan curve, while Radó's question is concerned only with the topological nature of a portion of the surface. However, a small modification of the method yields the answer to this question also.

§§1–4 of the paper are concerned with the properties of excrescences and the possibility of their removal. §§5–7 contain lemmas needed for the proof in §8 of the principal theorem, Theorem 8.1. In §8 we also establish three corollaries to Theorem 8.1, and prove Theorem 8.2, which answers Radó's question.

Notation and Definitions. The area of a surface will always be understood in the sense of Lebesgue, and will be denoted by $L(S)$. We shall also speak of the areas of surfaces $x = x(u, v)$, etc., where (u, v) ranges over a doubly connected region; the usual discussions of area apply equally well to such surfaces. The portion of S corresponding to a subset B' of the region of variability B of (u, v) will be denoted by $S[B']$. If B' be a connected open set we can define $L(S[B'])$, for example, as the least upper bound of $L(S[B''])$ for all *Jordan* regions B'' contained in B'; this quantity is clearly invariant under topological mappings of B'. We shall assume throughout that all areas mentioned are finite.

For any point set M in the (u, v)-plane we shall denote the boundary of M by M^* by and the set $M + M^*$ by $\bar{M}$. (But the symbol $\bar{S}$ will *not* denote the "closure" of the surface S.)

The letters E, F, G will have the meanings usual in differential geometry. The statement that the surface

$$S\colon\ x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v), \qquad (u, v) \text{ on } \bar{B}$$

is bounded by a curve Γ means that as (u, v) transverses B^*, these equations yield a representation of Γ. Even when Γ is a *Jordan* curve this representation need not be one-to-one; in the terminology of Radó,[3] it need only be a continuous monotonic transformation of B^* into Γ.

---

[3] Radó (2), loc. cit.[1]; in particular, p. 872.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

ON THE ANALYTIC NATURE OF SURFACES OF LEAST AREA

BY E. J. MCSHANE

(Received December 4, 1933)

Introduction

There is an intimate connection between the problem of finding a minimal surface bounded by a given non-self-intersecting curve and the problem of finding a surface bounded by that curve and having least possible area. The very name "minimal surface" arises from the fact that the differential equations of a minimal surface are the Euler-Lagrange equations of the problem of least area, expressing the vanishing of the first variation of the area. Moreover, if the given curve bounds a surface of finite area (which is the only case for which the problem of least area has a meaning) the two problems have common solutions.[1]

Nevertheless, it is well known that a minimal surface need not have least area for its boundary, and also that a surface of least area need not be minimal. The following problem therefore suggests itself: Given a surface S, defined by equations

$$x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v) \qquad u^2 + v^2 \leqq 1,$$

and given a region B in the circle $u^2 + v^2 \leqq 1$. What conditions on x, y, z enable us to state that the portion of S corresponding to B is minimal?

The most general result at present in the literature is as follows: The portion of S corresponding to B is minimal if it projects orthogonally in a one-to-one way on some plane.[2] Radó asks if it be sufficient to assume that the equations $x = x(u, v)$, etc., carry B in a one-to-one way into a portion of S. One of the results of the present paper is to answer this question affirmatively.

Let us give the name "excrescence" to a portion of S, not itself a single point, which joins the rest of S in a single point (for example, a line-segment forming a spine projecting from S). Analytically, this means that there is an open subset B of the unit circle such that x, y, z are constant on the boundary of B without being constant on all of B. If S has excrescences it is clearly not minimal. But it would seem to be (and actually is) possible to remove all excrescences from S,

---

[1] Jesse Douglas, *Solution of the problem of Plateau*, Trans. Am. Math. Soc. **33** (1931), pp. 263–321. E. J. McShane, *Parameterizations of saddle surfaces, etc.*, Trans. Am. Math. Soc. **35** (1933) pp. 716–733. Tibor Radó: (1) *The problem of least area and the problem of Plateau*, Math. Zeitschrift, **32** (1930) pp. 763–796. (2) *An iterative process in the problem of Plateau*, Trans. Am. Math. Soc. **35** (1933) pp. 869–887.

[2] E. J. McShane, *On a certain inequality of Steiner*, Annals of Math., **33** (1932) pp. 125–138.

thus obtaining a surface $\bar{S}$ free from excrescences. We might then hope to prove that if S has least area for its boundary, the surface $\bar{S}$ obtained by removing excrescences from S is minimal. This is however false; the surface $x(u, v) = u, y(u, v) \equiv z(u, v) \equiv 0, 0 \leqq u, v \leqq 1$ has no excrescences and is not minimal. But the boundary of this surface is not a Jordan curve. We may therefore try to prove that if S is bounded by a Jordan curve Γ and has least area among all surfaces bounded by Γ, then the surface $\bar{S}$ is minimal. The proof of this statement is the principal object of the present paper. The result may also be stated thus: For a given Jordan curve Γ, the most general surface of least area bounded by Γ is a minimal surface plus excrescences. It follows that a surface of least area with a Jordan boundary and without excrescences is minimal.

This does not answer Radó's question, for it requires that the boundary Γ of the whole surface S be a Jordan curve, while Radó's question is concerned only with the topological nature of a portion of the surface. However, a small modification of the method yields the answer to this question also.

§§1–4 of the paper are concerned with the properties of excrescences and the possibility of their removal. §§5–7 contain lemmas needed for the proof in §8 of the principal theorem, Theorem 8.1. In §8 we also establish three corollaries to Theorem 8.1, and prove Theorem 8.2, which answers Radó's question.

Notation and Definitions. The area of a surface will always be understood in the sense of Lebesgue, and will be denoted by $L(S)$. We shall also speak of the areas of surfaces $x = x(u, v)$, etc., where (u, v) ranges over a doubly connected region; the usual discussions of area apply equally well to such surfaces. The portion of S corresponding to a subset B' of the region of variability B of (u, v) will be denoted by $S[B']$. If B' be a connected open set we can define $L(S[B'])$, for example, as the least upper bound of $L(S[B''])$ for all *Jordan* regions B'' contained in B'; this quantity is clearly invariant under topological mappings of B'. We shall assume throughout that all areas mentioned are finite.

For any point set M in the (u, v)-plane we shall denote the boundary of M by M^* by and the set $M + M^*$ by $\bar{M}$. (But the symbol $\bar{S}$ will *not* denote the "closure" of the surface S.)

The letters E, F, G will have the meanings usual in differential geometry. The statement that the surface

$$S: x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v), \qquad (u, v) \text{ on } \bar{B}$$

is bounded by a curve Γ means that as (u, v) transverses B^*, these equations yield a representation of Γ. Even when Γ is a *Jordan* curve this representation need not be one-to-one; in the terminology of Radó,[3] it need only be a continuous monotonic transformation of B^* into Γ.

---

[3] Radó (2), loc. cit.[1]; in particular, p. 872.

§1. Definition of Excrescences

Let the equations

$$(1.1) \qquad x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v), \qquad (u, v) \text{ on } \bar{B}$$

be a representation of a continuous surface S, the region $\bar{B}$ consisting of a Jordan curve B^* plus its interior B, and let B' be a connected open subset of B. We shall say that the equations

$$(1.2) \qquad x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v), \qquad (u, v) \text{ on } \bar{B}'$$

define an *excrescence* on the surface S provided that the functions $x(u, v)$, $y(u, v)$, $z(u, v)$ are all constant on the boundary of B', and are *not* all constant on B' itself. If equations (1.2) define an excrescence, we shall say that B' defines an excrescence, or is the *defining region* of an excrescence, and we shall say that $S[B']$ (or $S[\bar{B}']$) is an excrescence.

The open set B' need not be simply connected. Nevertheless, it is possible to prove that

(1.3) *If an open set B' defines an excrescence, then there is a connected and simply connected open set B'' defining an excrescence and containing B'.*

In order to construct the set B'', we first observe that the set of all simply closed polygons lying in B' and having rational coordinates for all vertices is a denumerable set. We denote these polygons by $\Pi_1^*, \Pi_2^*, \cdots$, and their interiors by $\Pi_1, \Pi_2, \cdots$. The set B'' is then defined to be the sum $\Pi_1 + \Pi_2 + \cdots$. We state five properties of the set B'':

(1.4) *B'' is open;* for it is a sum of open sets.

(1.5) *B'' contains B'.* For every point P of B' can be enclosed in a square lying in B' and having rational vertices, and this square is one of the Π_n^*.

(1.6) *B'' is connected.* Let P_1, P_2 be points of B'' belonging to Π_m, Π_n respectively, and let P_1', P_2' be points of the polygons Π_m^*, Π_n^* respectively. The points P_1', P_2' belong to B' and can be connected in B', hence in B''. Also P_1, P_1' can be connected in $\bar{\Pi}_m$, hence in B'', and likewise P_2, P_2'. Hence P_1, P_2 can be connected in B''.

(1.7) *B'' is simply connected.* Let J be any Jordan curve in B''; we are to show that the interior of J also belongs to B''. Each point of J belongs to a region Π_n, hence a finite number of the Π_n cover J. The sum of this finite number of regions Π_n we call M. Let J_1 be the outer boundary of $\bar{M}$ (i.e., the boundary points accessible from infinity). J_1 is a simply closed polygon in B' with rational vertices, hence is one of the Π_n^*, say Π_q^*. The region Π_q contains J and is simply connected, hence Π_q (and therefore B'') contains the interior of J.

(1.8) *The boundary of B'' is contained in the boundary of B'.* Let P be a boundary point of B''; by (1.4) and (1.5), P does not belong to B'. In every square neighborhood N of P there is a point P_1 of B'', contained in a set Π_n. Since P is not in Π_n, the segment PP_1 meets the polygon Π_n^* in a point P_2, which belongs to B' and is in N. Hence P is a limit point of B'.

It follows at once that the region B'' satisfies (1.3). For x, y, z are constant on the boundary of B', hence by (1.8) they are constant on the boundary of B'', and x, y, z are not all constant on B', so by (1.5) they are not all constant on B''.

§2. Complete Excrescences

It is of course possible that if an open set B' defines an excrescence on S, it is contained in a larger set defining an excrescence. This suggests searching for the largest set containing B' and defining an excrescence.

Let R be an arbitrarily chosen point of B', and let B_0 be the set of all points Q of B (cf. equation (1.1)) which are contained in an open set $B(Q)$ containing R and defining an excrescence. By (1.3), we may assume that $B(Q)$ is simply connected. We state that B_0 defines an excrescence. First, B_0 is open; for if Q is a point of B_0, it is contained in an open set $B(Q)$ with the above properties, and every point in a neighborhood of Q belongs to $B(Q)$, hence to B_0. Second, the functions $x(u, v)$, $y(u, v)$, $z(u, v)$ are all constant on the boundary B_0^* of B_0. For let P_1, P_2 be any two points of B_0^*, and let N_1, N_2 be square neighborhoods of P_1, P_2 respectively on which the oscillations of x, y, z are all less than ϵ (ϵ an arbitrary positive number). The neighborhoods N_i $(i = 1, 2)$ contain points Q_i which belong respectively to sets $B(Q_i)$ defining excrescences and containing R. Hence they also contain boundary points of $B(Q_i)$. Let x_i, y_i, z_i be the constant values of x, y, z on the boundaries of $B(Q_i)$. It is readily seen that

$$|x(P_i) - x_i| < \epsilon, \qquad (i = 1, 2), \tag{2.1}$$

with like inequalities for y, z.

The sets $B(Q_i)$ have the point R in common, and their boundaries are continua. Hence either one of them contains the other or else they have a boundary point in common. In the latter case $x_1 = x_2$, $y_1 = y_2$, $z_1 = z_2$, so that (2.1) implies

$$|x(P_1) - x(P_2)| < 2\epsilon. \tag{2.2}$$

In the former case, suppose $B(Q_1) \subset B(Q_2)$. Then Q_1 is a point of $B(Q_2)$, and N_1 contains a boundary point of Q_2, so that

$$|x(P_1) - x_2| < \epsilon. \tag{2.3}$$

This with (2.1) implies (2.2). Since inequality (2.2) holds for every $\epsilon > 0$, we have

$$x(P_1) = x(P_2), \tag{2.4}$$

with like equations for y, z. Hence x, y, z are all constant on the boundary B_0^* of B_0.

Since B_0 contains B', the functions x, y, z are not all constant on B_0. Hence B_0 defines an excrescence.

It is easily seen that we would have been led to the same set B_0 had we chosen any other point R' of B_0 instead of R.

Suppose that B_2 is an open set defining an excrescence, and that B_2 contains a point R' of B_0. By the definition of B_0, the whole set B_2 is contained in B_0. Hence we are justified in calling the portion of S corresponding to B_0 a *complete* excrescence.

The set B_0 is simply connected. For by (1.3) B_0 is contained in a simply connected region B'' defining an excrescence. But by the preceding paragraph $B'' \subset B_0$; hence $B'' = B_0$, and B_0 is simply connected.

If the open sets B_0, B_1 define two complete excrescences, they are non-overlapping. Hence *a continuous surface can have at most a denumerable set of complete excrescences.*

§3. Removal of an Excrescence

Let the surface S be defined by equations (1.1), and let the simply-connected open set B_1 define an excrescence. We denote by x_1, y_1, z_1 the constant values of x, y, z on the boundary of B_1. The continuous surface S_1 defined by the equations

$$(3.1)\qquad \begin{aligned} &x = x(u, v), \quad y = y(u, v), \quad z = z(u, v), \quad (u, v) \text{ on } \bar{B} - B_1, \\ &x = x_1, \quad y = y_1, \quad z = z_1, \quad (u, v) \text{ on } B_1 \end{aligned}$$

will be said to be obtained from S by removal of the excrescence $S[B_1]$. We wish to show that the area of S_1 is equal to the area of S minus that of the excrescence. This would be immediate if the boundary of B_1 were a Jordan curve; since this is not necessarily the case, we introduce a new parametric representation of S.

Let the functions $u = u(\rho, \vartheta)$, $v = v(\rho, \vartheta)$ map the interior of the unit circle $0 \leqq \rho < 1$ in a one-to-one and continuous way on the region B_1. We define

$$(3.2)\qquad \xi(\rho, \vartheta) = x(u(\rho, \vartheta), v(\rho, \vartheta)),$$

with analogous definitions for $\eta(\rho, \vartheta)$, $\zeta(\rho, \vartheta)$. It is readily seen that as $\rho \to 1$ we have $\xi \to x_1$, $\eta \to y_1$, $\zeta \to z_1$; so if we assign the functions ξ, η, ζ these values for $\rho = 1$, they become continuous, hence uniformly continuous, on the whole circle $0 \leqq \rho \leqq 1$. Now let $\varphi(\rho) = 2\rho$ if $0 \leqq \rho \leqq \frac{1}{2}$ and $\varphi(\rho) = 1$ if $\frac{1}{2} < \rho \leqq 1$, and define

$$(3.3)\qquad \bar{x}(u, v) \equiv \bar{x}(u(\rho, \vartheta), v(\rho, \vartheta)) \equiv \xi(\varphi(\rho), \vartheta), \qquad (u, v) \text{ on } B_1,$$

with analogous definitions for $\bar{y}(u, v)$ and $\bar{z}(u, v)$. On the portion of B_1 exterior to the Jordan curve $u = u(\frac{1}{2}, \vartheta)$, $v = v(\frac{1}{2}, \vartheta)$ the functions $\bar{x}, \bar{y}, \bar{z}$ have the constant values x_1, y_1, z_1 respectively.

The surface $\bar{S}$ defined by the equations

$$(3.4)\qquad \begin{aligned} &x = x(u, v), \quad y = y(u, v), \quad z = z(u, v), \quad (u, v) \text{ on } \bar{B} - B_1 \\ &x = \bar{x}(u, v), \quad y = \bar{y}(u, v), \quad z = \bar{z}(u, v), \quad (u, v) \text{ on } B_1 \end{aligned}$$

is easily seen to be continuous; we state that it is identical with S. For let $\varphi_n(\rho)$ be a piece-wise linear function with the corner-values

$$\varphi_n(0) = 0, \quad \varphi_n(\tfrac{1}{2}) = 1 - 2/n, \quad \varphi_n(1 - 1/n) = 1 - 1/n, \quad \varphi_n(1) = 1 \ (n > 2).$$

The equations

$$\begin{aligned} &\bar{u} = u, \qquad \bar{v} = v, \qquad (u, v) \text{ on } \bar{B} - B_1 \\ &\bar{u} = u(\varphi_n(\rho), \vartheta), \qquad v = v(\varphi_n(\rho), \vartheta), \qquad (u, v) \text{ on } B_1 \end{aligned} \tag{3.5}$$

map $\bar{B}$ topologically on itself, all points of B exterior to the curve

$$u = u(1 - 1/n, \vartheta), \qquad v = v(1 - 1/n, \vartheta)$$

being left unchanged. Hence if we set

$$\begin{aligned} &x = x_n(u, v) = x(u, v), \qquad (u, v) \text{ on } \bar{B} - B_1, \\ &x = x_n(u, v) = \xi(\varphi_n(\rho), \vartheta), \qquad (u, v) \text{ on } B_1, \end{aligned} \tag{3.6}$$

with analogous expressions for y_n, z_n, the equations

$$x = x_n(u, v), \qquad y = y_n(u, v), \qquad z = z_n(u, v), \qquad (u, v) \text{ on } \bar{B} \tag{3.7}$$

form another representation of S. But as n increases, the functions $x_n(u, v)$ are unaltered on $\bar{B} - B_1$, and because of the uniform continuity of $\xi(\rho, \vartheta)$ and the uniform approach $\varphi_n \to \varphi$ the relation $\lim x_n(u, v) = \bar{x}(u, v)$ holds uniformly on B_1. Hence $\lim x_n(u, v) = \bar{x}(u, v)$ uniformly on $\bar{B}$, like relations holding for y_n and z_n. It follows that $\lim S = \bar{S}$, so that S and $\bar{S}$ are the same surface.

Let us denote by B_2 the interior of the Jordan curve $u = u(\frac{1}{2}, \vartheta), v = v(\frac{1}{2}, \vartheta)$. Recalling (3.1) and (3.4), we see that $\bar{S}[\bar{B} - B_2] = S_1[\bar{B} - B_2]$, while $S_1[B_2]$ is the single point (x_1, y_1, z_1) and $\bar{S}[B_2]$ is the excrescence. The boundary of B_2 is a Jordan curve whose image on $\bar{S}$ is rectifiable, being in fact the single point (x_1, y_1, z_1); hence

$$\begin{aligned} L(\bar{S}) &= L(\bar{S}[\bar{B} - B_2]) + L(\bar{S}[B_2]), \\ L(S_1) &= L(S_1[\bar{B} - B_2]) + L(S_1[B_2]) = L(S_1[\bar{B} - B_2]). \end{aligned}$$

This yields

$$L(S_1) = L(\bar{S}) - L(\bar{S}[B_2]) = L(S) - L(\bar{S}[B_2]); \tag{3.8}$$

that is, the area of the surface obtained by removing from S the excrescence $\bar{S}[B_2]$ or $S[B_1]$ is the area of S minus that of the excrescence.

§4. Removal of All Excrescences

Let the surface S be defined by equations (1.1), and let $S[B_1], S[B_2], \cdots$ be all of its complete excrescences. From S we obtain

$$\bar{S}\text{:}\ x = \bar{x}(u, v), \qquad y = \bar{y}(u, v), \qquad z = \bar{z}(u, v), \qquad (u, v) \text{ on } \bar{B} \tag{4.1}$$

by setting $\bar{x} = x$, $\bar{y} = y$, $\bar{z} = z$ on $\bar{B} - \Sigma B_i$, and on each B_i giving x, y, z the constant values assumed by x, y, z on the boundary B_i^* of B_i. The surface $\bar{S}$ is continuous; we call it the surface obtained from S by removal of all excrescences. Concerning $\bar{S}$ we make three statements:

(4.2) The boundary curve of $\bar{S}$ is the same as that of S.

(4.3) $\bar{S}$ has no excrescences.

$$L(\bar{S}) = L(S) - \sum_{i=1}^{\infty} L(S\,[B_i]). \tag{4.4}$$

Statement (4.2) is obvious, since $x = \bar{x}$, $y = \bar{y}$, $z = \bar{z}$ on the boundary of B. Suppose that $\bar{S}$ has an excrescence; it then has a complete excrescence, $\bar{S}\,[B_0]$. No point of the boundary B_0^* of B_0 can belong to any B_i; for if a point P of B_0^* belongs to B_i, then B_i and B_0 have a common interior point R. On the boundary B_i^* we have $x = \bar{x} = \text{const.}$, $y = \bar{y} = \text{const.}$, $z = \bar{z} = \text{const.}$, hence by the completeness of $\bar{S}\,[B_0]$ the whole region B_i belongs to B_0, and P is an interior point of B_0, which is impossible. Hence on B_0^* we have $x = \bar{x}$, etc. Either B_0 has no point in common with ΣB_i, or it has such a point. If it has none, then $x = \bar{x}$, etc., on B_0, and $S\,[B_0]$ is an excrescence, which is a contradiction. Suppose it has a point in common with B_i; since $S\,[B_i]$ is a complete excrescence and $x = \bar{x}$, etc., on B_0^*, the whole of B_0 is contained in B_i, and $\bar{x}$, $\bar{y}$, $\bar{z}$ are constant on B_0, which is a contradiction. Hence $\bar{S}$ has no excrescences, and (4.3) is established.

To prove (4.4), we first observe that if S' is obtained from S by assigning x', y', z' the constant boundary values of x, y, z on any finite or infinite subset of the B_i and setting $x' = x$, $y' = y$, $z' = z$ elsewhere, then the functions x', y', z' have the same modulus of continuity as x, y, z. In other words, all such functions x', y', z' are equi-continuous, no matter what subset of the B_i be chosen. Let us define

$$\begin{aligned} x_n(u, v) &= x(u, v), \qquad (u, v) \text{ on } \bar{B} - \sum_{i=1}^{n} B_i, \\ x_n(u, v) &= x_i, \qquad (u, v) \text{ on } B_i,\ i = 1, \cdots, n, \end{aligned} \tag{4.5}$$

where x_i is the constant value of x on the boundary B_i^*, and let us define y_n, z_n analogously. For the surface S_n defined by

$$x = x_n(u, v), \qquad y = y_n(u, v), \qquad z = z_n(u, v), \qquad (u, v) \text{ on } \bar{B} \tag{4.6}$$

we know by (3.8) that

$$L(S_n) = L(S) - \sum_{i=1}^{n} L(S\,[B_i]). \tag{4.7}$$

The functions x_n, y_n, z_n converge at each point to $\bar{x}$, $\bar{y}$, $\bar{z}$ respectively, and since they are equi-continuous the convergence is uniform. Hence $S_n \to \bar{S}$, and

$$L(\bar{S}) \leqq \liminf L(S_n) = L(S) - \sum_{i=1}^{\infty} L(S[B_i]). \tag{4.8}$$

On the other hand, let us define

$$\bar{x}_n(u, v) = x(u, v), \qquad (u, v) \text{ on } \bar{B} - \sum_{i=n+1}^{\infty} B_i,$$

$$\bar{x}_n(u, v) = \bar{x}(u, v), \qquad (u, v) \text{ on } \sum_{i=n+1}^{\infty} B_i,$$

and let us define $\bar{y}_n$, $\bar{z}_n$ analogously. The functions $\bar{x}_n$, etc. are equi-continuous, and converge at each point to $x(u, v)$; hence the convergence is uniform. Consequently the surface $\bar{S}_n$ defined by

$$x = \bar{x}_n(u, v), \qquad y = \bar{y}_n(u, v), \qquad z = \bar{z}_n(u, v), \qquad (u, v) \text{ on } \bar{B}$$

converges to the surface S.

But $\bar{S}$ is obtained from S_n by removing the excrescences

$$\bar{S}_n[B_i] = S[B_i], \qquad i = 1, 2, \cdots, n,$$

so that

$$L(\bar{S}) = L(\bar{S}_n) - \sum_{i=1}^{n} L(S[B_i]).$$

Hence

$$L(S) \leqq \liminf L(\bar{S}_n) = \lim \left\{ L(\bar{S}) + \sum_{i=1}^{n} L(S[B_i]) \right\} \tag{4.9}$$

$$= L(\bar{S}) + \sum_{i=1}^{\infty} L(S[B_i]).$$

Inequalities (4.8) and (4.9) together establish (4.4).

An immediate consequence of statements (4.2) and (4.4) is the following

(4.10) *If the surface S is bounded by a curve Γ and has the least area among all surfaces bounded by Γ, then S has no excrescence of positive area.*

For if it had such an excrescence, the surface $\bar{S}$ would have the same boundary and smaller area, which is impossible.

§5. A Lemma on Projections

From now on we find it more convenient to restrict ourselves to representations

$$(5.1) \qquad x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v), \qquad (u, v) \text{ on } Q,$$

where Q is the unit square $0 \leqq u \leqq 1, 0 \leqq v \leqq 1$. If equations (4.1) represent a surface $\bar{S}$ without excrescences, the parametrization can be changed to the form (4.1) by a continuous one-to-one mapping of $\bar{B}$ on Q, which clearly leaves $\bar{S}$ free of excrescences.

We now prove a lemma needed for the next section.

LEMMA 5.1. *If equations* (5.1) *represent a surface* S, *and* B *is an open set such that*[4] $z(u, v) = c = \text{const.}$ *on the boundary* B^* *of* B *and* $z(u, v) \neq c$ *on* B, *and if* $\bar{S}$ *is the surface*

$$(5.2) \qquad x = x(u, v), \qquad y = y(u, v), \qquad z = \bar{z}(u, v), \qquad (u, v) \text{ on } Q,$$

where $\bar{z} = c$ *on* B *and* $\bar{z} = z$ *elsewhere, then*

$$L(\bar{S}) \leqq L(S).$$

We suppose to be specific that $z > c$ on B.

Let $\{\Pi_n\}$ be a sequence of polyhedra tending to S such that

$$(5.3) \qquad \lim L(\Pi_n) = L(S).$$

For each Π_n we can find[5] a representation

$$(5.4) \qquad x = x_n(u, v), \qquad y = y_n(u, v), \qquad z = z_n(u, v), \qquad (u, v) \text{ on } Q$$

such that the functions x_n, etc., are linear on each of a set of triangles filling Q, and

$$(5.5) \qquad \begin{aligned} &\lim x_n(u, v) = x(u, v), \qquad \lim y_n(u, v) = y(u, v), \\ &\lim z_n(u, v) = z(u, v), \qquad \text{uniformly on } Q. \end{aligned}$$

There is clearly no loss of generality in assuming

$$(5.6) \qquad z_n(u, v) < z(u, v) \text{ for all } (u, v) \text{ on } Q.$$

The set of points for which $z_n > c$ consists of a finite number of triangles. These triangles can have no point in common with B^*, for on B^* we have

$$z_n < z = c.$$

Those of the triangles interior to B we add together to form the set R_n. We define $\bar{z}_n$ to be equal to c on R_n and to z_n elsewhere. The function $\bar{z}_n$ is continuous. Moreover, if we define the polyhedron $\overline{\Pi}_n$ by the equation

---

[4] Of course we could equally well permute x, y, z cyclically.

[5] E. J. McShane, *Semi-Continuity of Double Integrals in the Calculus of Variations*, Annals of Math., **33** (1932); in particular, pp. 468–469.

(5.7) $x = x_n(u, v), \qquad y = y_n(u, v), \qquad z = \bar{z}_n(u, v), \qquad (u, v) \text{ on } Q$

we have

(5.8) $$L(\bar{\Pi}_n) \leqq L(\Pi_n);$$

for the triangles of Π_n corresponding to R_n have been replaced by their orthogonal projections on the plane $z = c$.

The functions $\bar{z}_n(u, v)$ tend uniformly to the limit function $\bar{z}(u, v)$. For let ϵ be any positive number, and let n be large enough so that

(5.9) $$0 < z(u, v) - z_n(u, v) < \epsilon$$

for all (u, v) on Q. Then for every point not belonging to B we have

$$\bar{z}_n = z_n, \qquad \bar{z} = z,$$

so that

(5.10) $$0 = \bar{z}(u, v) - \bar{z}_n(u, v) < \epsilon.$$

For every point of R_n we have $\bar{z}_n = \bar{z} = c$, while for every point of $B - R_n$ we have $\bar{z}_n \leqq c$, so that

$$0 \leqq \bar{z} - \bar{z}_n < z - \bar{z}_n = z - z_n < \epsilon.$$

Hence the inequality $0 \leqq \bar{z} - \bar{z}_n < \epsilon$ holds for all (u, v). It follows that the polyhedra $\bar{\Pi}_n$ tend to the surface $\bar{S}$, and hence by (5.8) we find

$$L(\bar{S}) \leqq \liminf L(\bar{\Pi}_n) \leqq \lim L(\Pi_n) = L(S),$$

establishing the lemma.

Remark: In the proof we used only the relation $z \geqq c$ on B. Hence if B be any region such that $z = c$ on B^*, we can first apply lemma 5.1 to the subset B' of B on which $z > c$ and then to the subset B'' on which $z < c$, obtaining a surface S satisfying (5.2) and also having $L(\bar{S}) \leqq L(S)$.

§6. An Alteration of the Surface

In the next section we shall wish to consider the points (u, v) for which $z(u, v) > c$. This set (if not vacuous) is open, but may not be simply connected. We now establish a lemma by which simple connectedness can be gained without loss of any desirable property of the surface S.

LEMMA 6.1. *Let equations* (5.1) *represent a continuous surface* S, *and let* B *be a connected open subset of* Q *such that*

(a) $$z(u, v) = c = \text{const. on the boundary } B^* \text{ of } B;$$

(b) $$z(u, v) > c \text{ on } B.$$

Then there exists a surface

$$(6.1) \qquad \bar{S}\colon\ x = x(u, v), \qquad y = y(u, v), \qquad z = \bar{z}(u, v), \qquad (u, v) \text{ on } Q$$

with the following properties:

(c) $\bar{z}(u, v) = z(u, v)$ *for all* (u, v) *not in* B'', *where* B'' *is the simply connected open set formed from* B *as in* §1;

(d) *on* B'' *the inequality*

$$z(u, v) > c$$

holds;

(e) *the area of* $\bar{S}$ *does not exceed that of* S:

$$(6.2) \qquad L(\bar{S}) \leqq L(S).$$

Let the polygons Π_n^* be those discussed in §1. On Π_1^* we have $z > c$; since Π_1^* consists of a closed set of points, on it we have

$$(6.3) \qquad z(u, v) \geqq m_1 > c.$$

The set of points such that $z < m_1$ is open, and by (6.3) contains no point of Π_1^*. Let O_1 be the subset of this open set interior to Π_1^*. On the boundary of O_1 we have $z = m_1$. Hence if we define

$$(6.4) \qquad \begin{aligned} z_1(u, v) &= m_1 \text{ for } (u, v) \text{ on } O_1, \\ z_1(u, v) &= z(u, v) \text{ elsewhere,} \end{aligned}$$

and denote by S_1 the surface

$$x = x(u, v), \qquad y = y(u, v), \qquad z = z_1(u, v), \qquad (u, v) \text{ on } Q,$$

by lemma 5.1 we have

$$(6.5) \qquad L(S_1) \leqq L(S).$$

It is clear that

$$(6.6) \qquad z_1(u, v) \geqq m_1 > c \text{ on } \Pi_1$$

and that

$$(6.7) \qquad z_1(u, v) \geqq z(u, v) \text{ for all } (u, v) \text{ on } Q.$$

We define $z_2, z_3, \cdots$, by induction. Let m_n be the minimum of z_{n-1} on Π_n^*, and let O_n be the open set of points interior to Π_n^* on which $z_{n-1}(u, v) < m_n$. We define

$$(6.8) \qquad \begin{aligned} z_n(u, v) &= m_n, \qquad (u, v) \text{ on } O_n \\ z_n(u, v) &= z_{n-1}(u, v) \text{ elsewhere.} \end{aligned}$$

Then, defining S_n by the equations

$$x = x(u, v), \qquad y = y(u, v), \qquad z = z_n(u, v), \qquad (u, v) \text{ on } Q, \tag{6.9}$$

we have by lemma 5.1

$$L(S_n) \leqq L(S_{n-1}). \tag{6.10}$$

Also

$$z_n(u, v) \geqq m_n \text{ on } \Pi_n \tag{6.11}$$

and

$$z_n(u, v) \geqq z_{n-1}(u, v) \text{ for all } (u, v). \tag{6.12}$$

Since Π_n^* is in B we have

$$z(u, v) > c, \qquad (u, v) \text{ on } \Pi_n^*,$$

and applying (6.12) for $n = 1, 2, \cdots$, we find $z_{n-1}(u, v) > c$, (u, v) on Π_n^*. Hence

$$m_n > c. \tag{6.13}$$

It is easily seen that the modulus of continuity of $z(u, v)$ serves also for $z_1, z_2, \cdots$, so that all the $z_n(u, v)$ are equi-continuous. But by (6.12) the sequence $z_n(u, v)$ increases monotically, hence converges for each (u, v) in Q to a limit function $\bar{z}(u, v)$. The functions $z_n(u, v)$ being equi-continuous, the convergence is uniform. Hence

$$\lim z_n(u, v) = \bar{z}(u, v) \text{ uniformly on } Q. \tag{6.14}$$

Let (u_0, v_0) be any point of B''; it is then contained in one of the regions Π_n, say in Π_k. Then from (6.11) and (6.13) we have

$$z_k(u_0, v_0) \geqq m_k > c$$

and this, with (6.12) and (6.14) proves

$$\bar{z}(u_0, v_0) > c.$$

This establishes conclusion (d). Conclusion (c) is obvious, for a point (u, v) not belonging to B'' is not contained in any Π_n, and so

$$z(u, v) = z_1(u, v) = z_2(u, v) = \cdots = \bar{z}(u, v).$$

From (6.14) we see that $\lim S_n = \bar{S}$; hence, using (6.5) and (6.10),

$$L(\bar{S}) \leqq \liminf L(S_n) \leqq L(S),$$

proving conclusion (e).

§7. Monotonicity of the Defining Functions

Before proceeding to the proof of the principal lemma of this paper we introduce a definition. A function $f(u, v)$ is said to be *monotonic (in the sense of Lebesgue)* on a region B provided that for every open subset O with boundary O^* of B the inequality

$$(\min f(u, v) \text{ on } O^*) \leqq f(u, v) \leqq (\max f(u, v) \text{ on } O^*)$$

holds for all (u, v) in O. We now state

LEMMA 7.1. *Let the surface*

$$(7.1) \qquad S: x = x(u, v), \quad y = y(u, v), \quad z = z(u, v), \quad (u, v) \text{ on } Q$$

have least area for its boundary.[6] *Let B_0 be a Jordan region contained*[7] *in Q such that the portion $S[B_0]$ of the surface S corresponding to B_0 is free from excrescences. Then on B_0 the functions $x(u, v)$, $y(u, v)$, $z(u, v)$ are monotonic in the sense of Lebesgue.*

We prove the lemma indirectly. Suppose the lemma false; to be specific, suppose that $z(u, v)$ is not monotonic. There then exists an open subset B of B_0 such that $z = c = \text{const.}$ on the boundary B^* of B, while $z \neq c$ – say $z > c$ – on B. Let B'' be the simply connected region formed from B as in §1; then by lemma 6.1 there exists a surface

$$(7.2) \qquad \bar{S}: x = x(u, v), \quad y = y(u, v), \quad z = \bar{z}(u, v), \quad (u, v) \text{ on } Q$$

with the properties

$$(7.3) \qquad L(\bar{S}) \leqq L(S),$$

$$(7.4) \qquad \bar{z}(u, v) > c \text{ on } B'', \quad \bar{z}(u, v) = c \text{ on } B''^*,$$

$$(7.5) \qquad \bar{z}(u, v) = z(u, v) \text{ on } Q - B''.$$

From (7.5) it follows in particular that

$$(7.6) \qquad \bar{z} = z \text{ on } Q^* \text{ and on } B''^*,$$

when Q^*, B''^* are the boundaries of Q, B'' respectively.

The functions $x(u, v)$, $y(u, v)$ are not both constant on B''^*. For if they were $S[B'']$ would be an excrescence, contrary to hypothesis. Suppose then that $y(u, v)$ takes on distinct values y_0, y_1 at the points P_0'', P_1'', of B''^*; say $y_0 < y_1$. We can enclose P_0'', P_1'' in circles K_0, K_1 small enough so that for all points (u_0, v_0) of K_0 and all points (u_1, v_1) and K_1 we have

$$y(u_0, v_0) < y(u_1, v_1).$$

---

[6] That is, if we denote the boundary curve of S by Γ and let S' be any surface bounded by Γ, then $L(S) \leqq L(S')$. As always, $L(S)$ is assumed finite.

[7] B_0 may of course coincide with Q.

In K_0, K_1 respectively we choose points P_0', P_1' of B'', and we join these points by a polygonal path. We join P_0' to P_0'' by a line segment; this segment contains a point P_0 of B''^* which is nearest to P_0'. We define P_1 analogously. The points P_0, P_1 are then joined by a polygonal path J_0, which lies in B'' except for its ends and which we may assume to be non-self-intersecting. Also

$$y(P_0) < y(P_1). \tag{7.7}$$

Under the transformation

$$y = y(u, v), \qquad z = z(u, v) \tag{7.8}$$

the curve J_0 is mapped on a curve J_0' of the y z-plane. The point P_2 with coordinates $y = \frac{1}{2}(y(P_0) + y(P_1))$, $z = c$ is not a point of J_0'; for at every point except the end points of J_0' we have $z > c$, while at the end points we have $y = y(P_0)$, $y = y(P_1)$ respectively, so that by (7.7) P_2 is not an end point of J_0'. It is therefore possible to describe a (closed) circular region K_2 with center at P_2 having no point in common with J_0'. We now construct a (closed) circular region K_3 interior to K_2 and lying above the line $z = c$. Let $c + 2\epsilon$ be the least of the z-coordinates of all points of K_3.

Since $\epsilon > 0$, there exists a $\delta > 0$ such that for every point (u, v) whose distance from B''^* is less than δ, the inequality

$$z(u, v) < c + \epsilon$$

holds. We now construct (as is easily possible) a non-self-intersecting polygon J_1 lying in B'' and having the following properties:

(a) The initial point A_1 of J_1 lies on the segment P_1P_1' of J_0.

(b) The end-point A_0 of J_1 lies on the segment P_0P_0' of J_0.

(c) Every point of J_1 has distance less than δ from B''^*.

(d) Only the points A_1, A_0 of J_1 lie on J_0.

The curve J^* formed by traversing J_0 from A_0 to A_1 and then J_1 from A_1 to A_0 is a simply closed polygon. The set consisting of J^* plus its interior we call J.

If we denote the minimum and maximum of y on the circle K_3 by m_3, M_3 respectively, we have by the definition of K_3

$$y(P_0) < m_3 < M_3 < y(P_1).$$

Since $y(u, v)$ is continuous, it follows that if A_0, A_1 are sufficiently close to P_0, P_1 respectively, (as can be secured by choosing δ small enough) we have

$$y(A_0) < m_3 < M_3 < y(A_1).$$

Now let J' be the image of J^* under the transformation (7.8), and let D be any point of the circle K_3. As a variable point P traverses the curve J^* once in the positive sense, its image P' traverses the curve J' once. We state that the line DP' then makes a net revolution of -2π radians. For as P goes from A_0 to A_1 along J_0, the line DP' passes from an initial position in the second or third quadrant to a position in the first or fourth quadrant, but never assumes the

direction of the negative z-axis. As P continues from A_1 back to A_0 on the arc J_1, the line DP' returns to its original position; but the image of J_1 lies entirely below D, so DP' never assumes the direction of the positive z-axis. In other words,

(7.9) Every point of the circle K_3 has index -1 with respect to the image J' of the Jordan polygon J.

We now form a sequence $\{\Pi_n\}$ of polyhedra

$$x = x_n(u, v), \qquad y = y_n(u, v), \qquad z = z_n(u, v), \qquad (u, v) \text{ on } Q$$

such that[8]

(7.10) the functions x_n, etc., are quasi-linear; that is, x_n, etc., are linear on each of a set of triangles filling Q, and are continuous on Q;

$$\begin{aligned} &\lim x_n(u, v) = x(u, v), \qquad \lim y_n(u, v) = y(u, v), \\ &\lim z_n(u, v) = \bar{z}(u, v), \qquad \text{uniformly on } Q; \end{aligned} \tag{7.11}$$

$$\lim L(\Pi_n) = L(\bar{S}). \tag{7.12}$$

From condition (7.11) we see readily that we can also suppose

$$z_n(u, v) < \bar{z}(u, v), \qquad (u, v) \text{ on } Q. \tag{7.13}$$

Let us now define $\zeta(u, v)$ by the equations

$$\begin{aligned} \zeta(u, v) &= c, && (u, v) \text{ on } B'', \\ \zeta(u, v) &= \bar{z}(u, v), && (u, v) \text{ on } Q - B''. \end{aligned} \tag{7.14}$$

The set of points (u, v) such that $z_n(u, v) > c$ is an open set which by (7.4) and (7.13) has no point in common with the boundary of B''. The portion of this set contained in B'' we call R_n. We then have

$$\begin{aligned} &z_n > c \text{ on } R_n, \\ &z_n = c \text{ on the boundary } R_n^* \text{ of } R_n. \end{aligned} \tag{7.15}$$

Let us define

$$\begin{aligned} \bar{z}_n(u, v) &= c \qquad (u, v) \text{ on } R_n \\ \bar{z}_n(u, v) &= z_n(u, v), \qquad (u, v) \text{ on } Q - R_n. \end{aligned} \tag{7.16}$$

As in §5 we find that

$$\lim \bar{z}_n(u, v) = \zeta(u, v) \tag{7.17}$$

---

[8] Loc. cit.[5]

uniformly on Q. Consequently the polyhedra $\bar{\Pi}_n$ defined by

(7.18) $\bar{\Pi}_n$: $x = x_n(u, v), \quad y = y_n(u, v), \quad z = \bar{z}_n(u, v), \quad (u, v)$ on Q

tend to the surface Σ defined by the equations

(7.19) $x = x(u, v), \quad y = y(u, v), \quad z = \zeta(u, v), \quad (u, v)$ on Q,

and therefore

(7.20) $L(\Sigma) \leqq \liminf L(\bar{\Pi}_n)$.

The closed region $\bar{J}$ is interior to B'', hence on it we have $\bar{z} > c$, so that $\bar{z} \geqq c + \gamma > c$. By (7.11) this implies

(7.21) $J \subset R_n$ for all sufficiently large n.

As n increases, the image J_n' of the curve J^* under the transformation

(7.22) $y = y_n(u, v), \quad z = z_n(u, v)$

tends to the curve J', by (7.11). Hence for all sufficiently large n, the index of every point of the circle K_3 with respect to J_n' has the value -1, because of (7.9). It follows at once that

(7.23) *for all sufficiently large n, the projection*

$$y = y_n(u, v), \quad z = z_n(u, v), \quad (u, v) \text{ on } \bar{J}$$

of the surface

$$x = x_n(u, v), \quad y = y_n(u, v), \quad z = z_n(u, v), \quad (u, v) \text{ on } \bar{J}$$

on the yz-plane contains the circle K_3. The area of this projection being equal to $\iint_J |X_n|\, du\, dv$, we have

(7.24) $$\iint_J |X_n|\, du\, dv \geqq a_3$$

for all sufficiently large n, where a_3 is the area of K_3. On the other hand, for all large n we have $\bar{z}_n(u, v) = c$ on J, so that the jacobians $\bar{X}_n$, $\bar{Y}_n$ vanish identically on J:

(7.25) $\bar{X}_n(u, v) = \bar{Y}_n(u, v) = 0, \quad \bar{Z}_n(u, v) \equiv Z_n(u, v), \quad (u, v)$ on J.

Finally, by (7.12) the areas $L(\Pi_n)$ are bounded, say less than H, so that

(7.26) $$\iint_Q [X_n^2 + Y_n^2 + Z_n^2]^{\frac{1}{2}}\, du\, dv < H$$

for all n.

Let M_n be the subset of J on which

(7.27) $$|X_n(u, v)| > (a_3/2H)\,[X_n^2 + Y_n^2 + Z_n^2]^{\frac{1}{2}}.$$

By (7.26),

$$\int\int_{J-M_n} |X_n|\,du\,dv \leqq (a_3/2H)\cdot\int\int_{J-M_n}[X_n^2+Y_n^2+Z_n^2]^{\frac{1}{2}}\,du\,dv$$
$$< (a_3/2H)\cdot H = a_3/2;$$

hence, recalling (7.24),

$$\int\int_{M_n} |X_n|\,du\,dv \geqq a_3/2. \tag{7.28}$$

Since the functions x_n, y_n, z_n, $\bar{z}_n$ are all quasi-linear, the areas of the surfaces Π_n, $\overline{\Pi}_n$ are given by the classical double integral. Hence for all large n, we find from (7.25), (7.27) and (7.28)

$$\begin{aligned} L(\Pi_n) - L(\overline{\Pi}_n) &= \int\int_Q \{[X_n^2+Y_n^2+Z_n^2]^{\frac{1}{2}} - [\bar{X}_n^2+\bar{Y}_n^2+\bar{Z}_n^2]^{\frac{1}{2}}\}\,du\,dv \\ &\geqq \int\int_J \{[X_n^2+Y_n^2+Z_n^2]^{\frac{1}{2}} - |Z_n|\}\,du\,dv \\ &= \int\int_J \frac{X_n^2+Y_n^2}{[X_n^2+Y_n^2+Z_n^2]^{\frac{1}{2}}+|Z_n|}\,du\,dv \\ &\geqq \int\int_{M_n} (X_n^2/2[X_n^2+Y_n^2+Z_n^2]^{\frac{1}{2}})\,du\,dv \\ &\geqq \int\int_{M_n} |X_n|\,(a_3/4H)\,du\,dv \geqq a_3^2/8H. \end{aligned}$$

(The division is permissible if the faces of Π_n are non-degenerate, which may be assumed without loss of generality.) Hence

$$\liminf L(\overline{\Pi}_n) \leqq \lim L(\Pi_n) - a_3^2/8H = L(\bar{S}) - a_3^2/8H. \tag{7.29}$$

Comparing (7.29), (7.20) and (7.3), we find

$$L(\Sigma) \leqq L(S) - a_3^2/8H. \tag{7.30}$$

But Σ has the same boundary as S, and S has the least area for that boundary. Hence inequality (7.30) expresses a contradiction, and the lemma is established.

§8. From lemma 7.1 we at once deduce our principal theorem on surfaces of least area. This is

THEOREM 8.1. *Let Γ be a Jordan curve in x y z-space bounding at least one surface of finite area, and let*

$$S:\ x = x(u,v), \qquad y = y(u,v), \qquad z = z(u,v), \qquad (u,v) \text{ on } \bar{B} \tag{8.1}$$

be a surface bounded by Γ *and having least area among all surfaces bounded*[9] *by* Γ. *Then the surface* $\bar{S}$ *obtained from* S *by removal of excrescences is a minimal surface in the sense of differential geometry; in other words,* S *is a minimal surface plus excrescences.*[10]

Let us change the range of (u, v) from $\bar{B}$ to Q by mapping $\bar{B}$ topologically on Q. The surface $\bar{S}$ then has a representation

$$\bar{S}\colon\ x = \bar{x}(u, v), \qquad y = \bar{y}(u, v), \qquad z = \bar{z}(u, v), \qquad (u, v) \text{ on } Q.$$

By lemma 7.1, the functions $\bar{x}$, etc., are monotonic; and by hypothesis the surface S, and with it $\bar{S}$, has finite area. The surface $\bar{S}$ therefore has a representation[11]

$$(8.2)\qquad \bar{S}\colon\ x = x_1(u, v), \qquad y = y_1(u, v), \qquad z = z_1(u, v), \qquad u^2 + v^2 \leqq 1$$

with the following properties:

(a) For almost all fixed values of u in the interval $-1 \leqq u \leqq 1$, the functions x_1, y_1, z_1 are absolutely continuous functions of v; and for almost all fixed values of v in the interval $-1 \leqq v \leqq 1$, the functions x_1, y_1, z_1 are absolutely continuous functions of u;

(b) the functions E_1 and G_1 are summable over the unit circle, E_1 and G_1 having the meanings usual in geometry;

(c) for almost all (u, v) the equations

$$(8.3)\qquad E_1 = G_1, \qquad F_1 = 0$$

hold.

We can now apply the proof on pp. 732–733 of my paper on saddle-surfaces,[12] beginning with equation (6.12), to show that $\bar{S}$ is a minimal surface and that equations (8.2) carry the circle $u^2 + v^2 = 1$ in a one-to-one way into the boundary curve Γ.

Three corollaries follow readily.

COROLLARY 1. *If the surface* S *satisfies the hypotheses of Theorem* 8.1 *and is free of excrescences, it is a minimal surface.*

For then S coincides with $\bar{S}$, and $\bar{S}$ is minimal.

COROLLARY 2. *If* S *satisfies the hypotheses of theorem* 8.1, *and equations* (8.1) *carry a Jordan sub-region* B' *of the region* B *in a one-to-one way into a portion of the surface* S, *then the portion of* S *corresponding to* B' *is minimal.*

No interior point of B' can belong to any excrescence $S[B_1]$ on the surface S. For if B_1 and B' had a common interior point, we could construct a Jordan

---

[9] We really need only the following somewhat weaker statement: If S_1 is a surface defined by functions $x = x_1(u, v)$, $y = y_1(u, v)$, $z = z_1(u, v)$, (u, v) on B, and the functions x_1, y_1, z_1 are identically equal to x, y, z respectively on the boundary B^*, then $L(S_1) \geqq L(S)$.

[10] Which by (4.10) are of area 0.

[11] E. J. McShane, loc. cit.[1]

[12] Loc. cit.[1]

region J belonging to both B' and B_1. Since the portion of S corresponding to J is a Jordan surface, its area is positive.[13] Hence $L(S[B_1])$ would be positive, contrary to (4.10). Consequently, when we remove the excrescences from S, the functions x, etc., are left unchanged on B', and so the portion of S corresponding to B' is also a portion of $\bar{S}$, which is minimal.

COROLLARY 3. *Let the surface S have least area ($< \infty$) for its boundary curve Γ (which we do not assume to be simply closed). Let the portion of S corresponding to the Jordan sub-region B' of B be free from excrescences and let the image of the boundary B'^* of B' be a rectifiable Jordan curve. Then the portion of S corresponding to B' is minimal.*

Let $S'[B']$ be a surface defined by functions $x = x'(u, v)$, $y = y'(u, v)$, $z = z'(u, v)$, (u, v) on $\bar{B}'$ which are identical with x, y, z respectively on B'^*. We define the surface S' by the relations

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v), \quad (u, v) \text{ on } \bar{B} - \bar{B}',$$

$$x = x'(u, v), \quad y = y'(u, v), \quad z = z'(u, v), \quad (u, v) \text{ on } \bar{B}'.$$

Since $S[\bar{B} - B']$, $S[\bar{B}']$, $S'[\bar{B}']$ join along a rectifiable curve,

$$L(S) = L(S[\bar{B} - B']) + L(S[\bar{B}']),$$

$$L(S') = L(S[\bar{B} - B']) + L(S'[\bar{B}']).$$

From the least area property of S, we have

$$L(S[\bar{B}']) \leqq L(S'[\bar{B}']).$$

By Theorem 8.1 and footnote 10, this implies that $S[\bar{B}']$ is minimal.

REMARK. Corollary 3 clearly remains valid if the region B be multiply connected.

The hypothesis in Corollary 2 that S has a Jordan curve for its boundary is in fact unnecessary; we can prove

THEOREM 8.2. *Let the surface S have a finite area which is least among all surfaces with its boundary. Let the equations* (8.1) *carry a Jordan sub-region B' of B in a one-to-one way into a portion of S. Then the portion of S corresponding to B' is minimal in the sense of differential geometry.*

Let q be an arbitrary (closed) square region interior to B'. We wish to show that the portion of S corresponding to q is minimal. Equations (8.1) carry the boundary B'^* and the square q into two non-overlapping closed point sets in $x\,y\,z$-space. These sets have a positive distance ϵ.

We can assume without loss of generality that the region B of (8.1) is the unit square Q. As in the proof of corollary 2, the portion of S corresponding to B' is free from excrescences; hence by lemma 7.1 the functions $x(u, v)$, etc., are monotonic on B'. The portion of S corresponding to B' has finite area and is bounded

[13] For the boundary of $S[J]$ spans a minimal surface S'' such that $L(S'') \leqq L(S[J])$. But since S'' is minimal and is not a single point we have $L(S'') > 0$.

by a Jordan curve; hence it has a representation of the form (8.2), satisfying conditions (a), (b), and (c). It follows at once that there exist rectifiable curves on $S[B']$ separating $S[q]$ from the boundary of $S[B']$. Among these curves are Jordan curves. One way of proving this is to observe that among the curves is one of least length, which is necessarily non-self-intersecting. Another way is to notice that in the point set constituting such a curve we can construct[14] a Jordan curve separating $S[q]$ from the boundary, and this Jordan curve can be shown to be rectifiable. Corollary 3 can now be applied to show that $S[B']$, and with it $S[q]$, is minimal.

PRINCETON UNIVERSITY.

[14] R. L. Moore, *Concerning continuous curves in the plane*, Math. Zeitschrift, **15** (1922), pp. 254–260; also C. Kuratowski, *Sur les coupures du plan*, Fund. Math., **6** (1924), pp. 130–145.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

DERIVATIVES, DIFFERENCE QUOTIENTS AND TAYLOR'S FORMULA II[1]

BY HASSLER WHITNEY

(Received September 12, 1934)

1. **Introduction.** In Paper I[2] we gave conditions under which a function of one variable $f(x)$ is a polynomial of degree at most $m-1$, or has a continuous m^{th} derivative, in terms of the uniform convergence of the m^{th} difference quotient. Here we shall give similar results in two dimensions; the extension to n dimensions is immediate. The functions involved are always supposed defined in a certain (open) region R of space.

2. **Partial difference quotients.** Let $f(x, y)$ be defined in R. If i and j are two non-negative integers and h is $\neq 0$, the partial difference quotient $\Delta_h^{i\,j} f(x, y)$ is defined by the equation

$$\Delta_h^{i\,j} f(x, y) = \frac{1}{h^{i+j}} \sum_{k=0}^{i} \sum_{l=0}^{j} (-1)^{i+j-k-l} \binom{i}{k}\binom{j}{l} f(x + kh, y + lh). \tag{1}$$

Thus for instance

$$\Delta_h^{2\,0} f(x, y) = \frac{1}{h^2} [f(x + 2h, y) - 2f(x + h, y) + f(x, y)],$$

$$\Delta_h^{1\,1} f(x, y) = \frac{1}{h^2} [f(x + h, y + h) - f(x, y + h) - f(x + h, y) + f(x, y)].$$

It is easily seen that

$$\Delta_h^{i+1,\,j} f(x, y) = \frac{1}{h} [\Delta_h^{i\,j} f(x + h, y) - \Delta_h^{i\,j} f(x, y)], \tag{2}$$

with a similar equation for $\Delta_h^{i,\,j+1} f(x, y)$.

(a) Suppose the functions $f_{ij}(x, y)$ are defined in R for $i + j \leqq m$. Suppose

$$f_{00}(x + \xi, y + \eta) = \sum_{i+j \leqq m} \frac{f_{ij}(x, y)}{i!\, j!} \xi^i \eta^j + R(x, y; \xi, \eta), \tag{3}$$

where for each point (x^*, y^*) of R and each $\epsilon > 0$ there is a $\delta > 0$ such that if (x, y) and $(x + \xi, y + \eta)$ are any two points within δ of (x^*, y^*), then

$$| R(x, y; \xi, \eta) | \leqq (\xi^2 + \eta^2)^{m/2} \epsilon. \tag{4}$$

---

[1] Presented to the American Mathematical Society, December 27, 1933.

[2] Derivatives, difference quotients and Taylor's formula I, this Bulletin, vol. **40** (1934), pp. 89–94.

We find as in Paper I for $i + j \leqq m$

$$\Delta_h^{i\,j} f_{00}(x, y) = \frac{1}{h^{i+j}} \sum_{k,\,l} (-1)^{i+j-k-l} \binom{i}{k}\binom{j}{l} \left[\sum_{p+q \leqq m} \frac{f_{pq}(x, y)}{p!\ q!} (kh)^p (lh)^q + R(x, y; kh, lh) \right]$$

$$= f_{ij}(x, y) + \sum{}' \frac{f_{i+p,\,j+q}(x, y)}{(i+p)!\ (j+q)!} h^{p+q} \sum_{k,\,l} (-1)^{i+j-k-l} \binom{i}{k}\binom{j}{l} k^{i+p} l^{j+q} + \frac{1}{h^{i+j}} \sum_{k,\,l} (-1)^{i+j-k-l} \binom{i}{k}\binom{j}{l} R(x, y; kh, lh), \tag{5}$$

where in Σ' those and only those terms occur in which

$$p \geqq 0, \qquad q \geqq 0, \qquad 0 < p + q \leqq m - i - j.$$

(3) and (4) show that $f_{00}(x, y)$ is continuous. As before, we see in succession that $\Delta_h^{i\,j} f_{00}(x, y) \to f_{ij}(x, y)$ uniformly in the neighborhood of any point of R and $f_{ij}(x, y)$ is continuous ($i + j = m, m - 1, \cdots, 1$).

(b) Suppose $f(x, y) = f_{00}(x, y)$ and $\partial^{i+j} f(x, y)/\partial x^i\, \partial y^j$ exists and equals the continuous function $f_{ij}(x, y)$ ($i + j \leqq m$). Then (3) and (4) hold, by Taylor's theorem, and hence $\Delta_h^{i\,j} f(x, y) \to f_{ij}(x, y)$ uniformly ($i + j \leqq m$).

(c) If $f(x, y) = \Sigma_{i+j \leqq m} a_{ij} x^i y^j$ is a polynomial of degree at most m, then a Taylor's formula (3) holds with $f_{ij}(x, y) \equiv i!\, j!\, a_{ij} (i + j = m)$ and $R(x, y; \xi, \eta) \equiv 0$. Hence $\Delta_h^{i\,j} f(x, y) \equiv i!\, j!\, a_{ij} (i + j = m)$. If $f(x, y)$ is of degree at most $m - 1$, $\Delta_h^{i\,j} f(x, y) \equiv 0$ ($i + j = m$).

3. **Functions with given m^{th} partial derivatives.** Suppose the continuous functions $f_{ij}(x, y)$ are defined for $i + j = m$ throughout an open square S with center at (a, b). Under what conditions does there exist a function $f_{00}(x, y)$ in R such that

$$\frac{\partial^{i+j}}{\partial x^i\, \partial y^j} f_{00}(x, y) = f_{ij}(x, y) \qquad (i + j = m)? \tag{6}$$

The answer is given by the following lemma.

---

[3] We have here used the fact that

$$\sum_{k,\,l} (-1)^{i+j-k-l} \binom{i}{k}\binom{j}{l} k^p\, l^q = \sum_{k=0}^{i} (-1)^{i-k} \binom{i}{k} k^p \cdot \sum_{l=0}^{j} (-1)^{j-l} \binom{j}{l} l^q,$$

which equals zero if $p < i$ or $q < j$, and equals $i!\, j!$ if $p = i$ and $q = j$; we then changed p and q to $i + p$ and $j + q$ respectively.

LEMMA 1. *Let the $f_{ij}(x, y)$ be given as above. If for any k, l such that $k + l = m - 1$,*

$$\int_x^{x'} f_{k+1,\,l}(s, y)\,ds + \int_y^{y'} f_{k,\,l+1}(x', s)\,ds + \int_{x'}^{x} f_{k+1,\,l}(s, y')\,ds + \int_{y'}^{y} f_{k,\,l+1}(x, s)\,ds = 0, \tag{7}$$

provided the path of integration lies in S, then there exists a function $f_{00}(x, y)$ in S satisfying (6).[3a]

The necessity of the condition merely reflects the fact that $\partial^{k+l} f_{00}(x, y)/\partial x^k\,\partial y^l$ is a single valued function. We prove the sufficiency. For any k and l, $k + l < m$, we define the function f_{kl} along the line $y = b$ by the equation

$$f_{kl}(x, b) = \int_a^x \cdots \int_a^{s_3} \int_a^{s_2} f_{m-l,\,l}(s_1, b)\,ds_1\,ds_2 \cdots ds_{m-k-l}. \tag{8}$$

We now extend the definitions of these functions throughout S:

$$f_{kl}(x, y) = \sum_{p=0}^{m-k-l-1} \frac{1}{p!} f_{k,\,l+p}(x, b)(y - b)^p + \int_b^y \cdots \int_b^{s_3} \int_b^{s_2} f_{k,\,m-k}(x, s_1)\,ds_1\,ds_2 \cdots ds_{m-k-l}. \tag{9}$$

If we show that

$$\frac{\partial}{\partial x} f_{kl}(x, y) = f_{k+1,\,l}(x, y), \qquad \frac{\partial}{\partial y} f_{kl}(x, y) = f_{k,\,l+1}(x, y), \tag{10}$$

(6) will follow. (7) gives

$$\int_b^{s_2} [f_{k,\,m-k}(x', s_1) - f_{k,\,m-l}(x, s_1)]\,ds_1 = \int_x^{x'} [f_{k+1,\,m-k-1}(s, s_2) - f_{k+1,\,m-k-1}(s, b)]\,ds;$$

hence

$$\int_b^y \cdots \int_b^{s_3} \int_b^{s_2} [f_{k,\,m-k}(x', s_1) - f_{k,\,m-k}(x, s_1)]\,ds_1\,ds_2 \cdots ds_{m-k-l} = \int_x^{x'} \int_b^y \cdots \int_b^{s_2} [f_{k+1,\,m-k-1}(s, s_2) - f_{k+1,\,m-k-1}(s, b)]\,ds_2 \cdots ds_{m-k-l}\,ds.$$

[3a] In n dimensions we must assume such a formula holds where the contour is any rectangle with sides parallel to any two coordinate axes.

On dividing both sides by $x' - x$ and letting x' approach x, we find the derivative with respect to x of the integral in (9). Differentiating (8) gives

$$\frac{\partial}{\partial x} f_{kl}(x, b) = f_{k+1,\, l}(x, b).$$

Hence, differentiating (9) with respect to x,

$$\begin{aligned}\frac{\partial}{\partial x} f_{kl}(x, y) &= \sum_{p=0}^{m-k-l-1} \frac{1}{p!} f_{k+1,\, l+p}(x, b)(y - b)^p \\ &\quad + \int_b^y \cdots \int_b^{s_3} [f_{k+1,\, m-k-1}(x, s_2) - f_{k+1,\, m-k-1}(x, b)]\, ds_2 \cdots ds_{m-k-l} \\ &= \sum_{p=0}^{m-k-l-1} \frac{1}{p!} f_{k+1,\, l+p}(x, b)(y - b)^p \\ &\qquad - \frac{1}{(m-k-l-1)!} f_{k+1,\, m-k-1}(x, b)(y - b)^{m-k-l-1} \\ &\qquad + \int_b^y \cdots \int_b^{s_2} f_{k+1,\, m-(k+1)}(x, s_1)\, ds_1 \cdots ds_{m-(k+1)-l} \\ &= f_{k+1,\, l}(x, y).\end{aligned}$$

The other equation follows immediately on differentiating (9) with respect to y.

4. **Difference quotients and contour integrals.** We prove the following lemma.

LEMMA 2. *If $\Delta_h^{ij} f(x, y)$ approaches the continuous function $f_{ij}(x, y)$ uniformly in the neighborhood of any point of R for $i + j = m$, then (7) is satisfied in any square S lying wholly in R.*

We first derive an identity in difference quotients. Let x and y be fixed numbers, let p and q be fixed integers, take $h \neq 0$, and set $x^* = x + ph$, $y^* = y + qh$. Replace x by $x + kh$ in (2), and sum. This gives

$$h \sum_{k=0}^{p-1} \Delta_h^{i+1,\, j} f(x + kh, y) = \Delta_h^{ij} f(x^*, y) - \Delta_h^{ij} f(x, y).$$

Similarly we find

$$h \sum_{k=0}^{p-1} \Delta_h^{i+1,\, j} f(x + kh, y^*) = \Delta_h^{ij} f(x^*, y^*) - \Delta_h^{ij} f(x, y^*),$$

$$h \sum_{k=0}^{q-1} \Delta_h^{i,\, j+1} f(x, y + kh) = \Delta_h^{ij} f(x, y^*) - \Delta_h^{ij} f(x, y),$$

$$h \sum_{k=0}^{q-1} \Delta_h^{i,\, j+1} f(x^*, y + kh) = \Delta_h^{ij} f(x^*, y^*) - \Delta_h^{ij} f(x^*, y).$$

These equations give

$$h\sum_{k=0}^{p-1}\Delta_h^{i+1,\,j}f(x+kh,y)+h\sum_{k=0}^{q-1}\Delta_h^{i,\,j+1}f(x^*,y+kh)$$

$$-h\sum_{k=0}^{p-1}\Delta_h^{i+1,\,j}f(x+kh,y^*)-h\sum_{k=0}^{q-1}\Delta_h^{i,\,j+1}f(x,y+kh)=0. \tag{11}$$

The functions $f_{ij}(x, y)$ are uniformly continuous in S. Hence, given x, x', y, y', and an $\epsilon > 0$, we can take a $\delta > 0$ so that if $|h| < \delta$, $|y' - y^*| < \delta$, and $x + ph \leqq x' < x + (p+1)h$, then

$$\left|\int_x^{x'} f_{i+1,\,j}(s,y')\,ds - h\sum_{k=0}^{p-1} f_{i+1,\,j}(x+kh,y^*)\right| < \epsilon.$$

As $\Delta_h^{i+1,\,j}f(x, y) \to f_{i+1,\,j}(x, y)$ uniformly in $S(i + 1 + j = m)$, we can take $\delta' < \delta$ so small that if $|h| < \delta'$, then

$$\left|h\sum_{k=0}^{p-1} f_{i+1,\,j}(x+kh,y^*) - h\sum_{k=0}^{p-1}\Delta_h^{i+1,\,j}f(x+kh,y^*)\right| < hp\epsilon \leqq |x'-x|\,\epsilon.$$

We derive similar inequalities corresponding to the other three terms of (11). As ϵ is arbitrarily small, (11) together with these inequalities gives (7), as required.

5. **Difference quotients and polynomials.** Theorem I of Paper I generalizes to two (or n) dimensions as follows:

THEOREM I. *Let $f(x, y)$ be a function defined in the square S and measurable in each variable separately. A necessary and sufficient condition that $f(x, y)$ be a polynomial of degree at most $m - 1$ is that $\Delta_h^{m0}f(x, y) \to 0$ and $\Delta_h^{0m}f(x, y) \to 0$ uniformly in S.*

The necessity of the condition follows from 2(c); we must prove the sufficiency. Take a fixed y_0, and consider $f(x, y_0)$ as a function of x. $\Delta_h^{m0}f(x, y_0) = \Delta_h^m f(x, y_0) \to 0$ uniformly, and hence $f(x, y_0)$ is a polynomial of degree at most $m - 1$ in x, by Paper I, Theorem I. Similarly $f(x_0, y)$ is a polynomial of degree at most $m - 1$ in y; it follows that $f(x, y)$ is a polynomial of degree at most $m - 1$ in x and y.[4] This is true if S is replaced by a region R.

---

[4] Choose m numbers $y_0, \cdots, y_{m-1}$; then

$$f(x, y_k) = a_{k0} + a_{k1}x + \cdots + a_{k,\,m-1}x^{m-1} \qquad (k = 0, \cdots, m-1).$$

But also

$$f(x, y) = b_0(x) + b_1(x)y + \cdots + b_{m-1}(x)y^{m-1}.$$

Setting $y = y_k (k = 0, \cdots, m - 1)$ in this equation and using the first equations, we find that the b's are polynomials in x.

6. **Existence of derivatives.** We are now ready to prove the fundamental theorem:

THEOREM II. *Let $f(x, y)$ be a function defined in R and measurable in each variable separately. A necessary and sufficient condition that $\partial^{i+j} f(x, y)/\partial x^i \partial y^j$ exist and equal the continuous function $f_{ij}(x, y)(i + j = m)$ is that*

$$\Delta_h^{ij} f(x, y) \to f_{ij}(x, y)$$

uniformly in the neighborhood of any point of $R(i + j = m)$.

If $f(x, y)$ is continuous, then $f_{ij}(x, y)$ is continuous, by 2(a).

The necessity of the condition was proved in 2(b); we must prove the sufficiency. Consider any square S in R. By Lemmas 1 and 2 there is a function $g(x, y)$ in S such that $\partial^{i+j} g(x, y)/\partial x^i \partial y^j = f_{ij}(x, y)$ $(i + j = m)$; hence, by 2(b), $\Delta_h^{ij} g(x, y) \to f_{ij}(x, y)$ uniformly in S. Set $\phi(x, y) = f(x, y) - g(x, y)$; then $\Delta_h^{ij} \phi(x, y) \to 0$ uniformly in S, and $\phi(x, y)$ is a polynomial of degree at most $m - 1$, by Theorem I. Therefore $\partial^{i+j} f(x, y)/\partial x^i \partial y^j = \partial^{i+j} g(x, y)/\partial x^i \partial y^j = f_{ij}(x, y)$ in S, and hence in R.

THEOREM III. *A necessary and sufficient condition that $f(x, y) = f_{00}(x, y)$ have continuous m^{th} partial derivatives in R is that there exist functions $f_{ij}(x, y)$ in R $(i + j \leqq m)$ such that (3) and (4) hold. In this case, $\partial^{i+j} f(x, y)/\partial x^i \partial y^j = f_{ij}(x, y)$ $(i + j \leqq m)$.*

The necessity of the condition is a consequence of Taylor's formula. To prove the sufficiency, 2(a) shows that $\Delta_h^{ij} f(x, y) \to f_{ij}(x, y)$ uniformly $(i + j \leqq m)$; hence, by Theorem II, $\partial^{i+j} f(x, y)/\partial x^i \partial y^j$ exists and equals $f_{ij}(x, y)$ $(i + j \leqq m)$.

HARVARD UNIVERSITY.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

FUNCTIONS DIFFERENTIABLE ON THE BOUNDARIES OF REGIONS[1]

BY HASSLER WHITNEY

(Received September 12, 1934)

1. **Introduction.** Let the function $f(x_1, \cdots, x_n)$ be defined in the bounded region[2] R of n-space E, and suppose f has continuous m^{th} partial derivatives in R, i.e. f "is of class C^m" in R. If B is the boundary of R, how shall we decide whether f is of class C^m in $R + B$? If the derivatives of f take on boundary values on B, it would be natural to define the derivatives on B as the limit of their values in R. But it is easy to construct a region R and a function f such that the k^{th} partial derivatives of $f(0 < k \leqq m)$ are continuous in $R + B$, whereas at a certain boundary point P of B, f is not continuous;[3] it seems unreasonable in this case to say that f is of class C^m in $R + B$.

If it is possible to extend the definition of f throughout a region containing $R + B$ so that it has continuous m^{th} partial derivatives there, we may then surely say that f is of class C^m in $R + B$; this is the definition we shall use. We show in this note that, for certain regions, for a function to be of class C^m in the closed region, it is sufficient that the m^{th} partial derivatives be continuous on the boundary.

We shall use a one-dimensional notation, as in a paper by the author AE.[4] Thus $f_k(x) = f_{k_1 \cdots k_n}(x_1, \cdots, x_n)$, $(x' - x)^l = (x_1' - x_1)^{l_1} \cdots (x_n' - x_n)^{l_n}$, $l! = l_1! \cdots l_n!$, etc. We set $\sigma_k = k_1 + \cdots + k_n$. $r_{xx'}$ is the distance between the two points x and x'. The fundamental definition is: $f(x) = f_0(x)$ is of class C^m in A if functions $f_k(x)$ and $R_k(x'; x)$ $(\sigma_k \leqq m)$ exist in A such that

$$(1) \qquad f_k(x') = \sum_{\sigma_l \leqq m - \sigma_k} \frac{f_{k+l}(x)}{l!} (x' - x)^l + R_k(x'; x)$$

for each $k(\sigma_k \leqq m)$, and R_k has the following property. Given the point x^0 of A and an $\epsilon > 0$, there is a $\delta > 0$ such that if x and x' are any two points of A within δ of x^0, then

$$(2) \qquad |R_k(x'; x)| \leqq r_{xx'}^{m-\sigma_k} \epsilon.$$

---

[1] Presented to the American Mathematical Society, December 27, 1933.

[2] The restriction that R be bounded is made merely for simplicity; the theorem holds equally well without it.

[3] We give an example with $m = 1$. In polar coordinates, let S be the spiral $r = \theta^{-1}$ $(\theta \geqq 0)$. We let R be a narrow region about S, and set $f(x, y) = F(r, \theta) = \log \theta$ for those values of r which are approximately θ^{-1}. Differentiating, we find $\partial f/\partial x = -\theta^{-1} \sin \theta$ and $\partial f/\partial y = \theta^{-1} \cos \theta$ approximately in R. Hence $\partial f/\partial x$ and $\partial f/\partial y \to 0$ at the origin; but $f \to \infty$ there.

[4] Analytic extensions of differentiable functions defined in closed sets, Transactions of the American Mathematical Society, Vol. 36 (1934), pp. 63–89.

If $f(x)$ has continuous m^{th} partial derivatives in R, it is of class C^m in R, as we see by setting $f_{k_1 \cdots k_n} = \partial^{k_1 + \cdots + k_n} f / \partial x_1^{k_1} \cdots \partial x_n^{k_n}$ and using Taylor's formula. The converse also is true.

LEMMA 1. *If $f(x)$ is of class C^m in the closed set A, then its definition can be extended throughout E so it will be of class C^m there.*

This is proved in AE, Lemma 2.

2. **The property P.** We shall say a point set A has the property P if there is a number ω such that any two points x and y of A are joined by a curve C in A of length $L \leqq \omega\, r_{xy}$.

LEMMA 2. *If R has the property P, so has $R + B$.*

Let x and y be points of $R + B$, and suppose first that y is in B while x is in R. Let $y_1, y_2, \cdots$ be a sequence of points of R such that $r_{y_s y} < 2^{-s-2}\, r_{xy}$. As $r_{xy_1} < 3r_{xy}/2$, there is a curve C_0 joining x and y_1 of length $< 3\omega\, r_{xy}/2$. In general, $r_{y_s y_{s+1}} < 2^{-s-1}\, r_{xy}$, and hence there is a curve C_s joining y_s and y_{s+1} of length at most $2^{-s-1}\, \omega\, r_{xy}$. From the curves $C_0, C_1, C_2, \cdots$ we pick out a curve C joining x to y; its length is

$$L < 3\omega r_{xy}/2 + 2^{-2}\, \omega r_{xy} + 2^{-3}\, \omega r_{xy} + \cdots = 2\omega r_{xy}.$$

If both x and y are in B, we take a point x' with $r_{x'x} < r_{xy}/2$, find curves joining it to x and y, of lengths $L_1 < \omega r_{xy}$ and $L_2 < 3\omega r_{xy}$ respectively, as above; from these we pick out a curve C joining x and y, of length $L < 4\omega r_{xy}$. Thus the number $\omega' = 4\omega$ will do for $R + B$. Evidently any number $\omega' > \omega$ will do.

3. **A remainder formula.** We prove here that *if* (1) *holds on the rectifiable curve C of length L with end points x^* and x', then $R_k(x'; x^*)(\sigma_k \leqq m)$ is given by the Stieltjes integral*[4a]

$$R_k(x'; x^*) = - \sum_{\sigma_l = m - \sigma_k} \frac{1}{l!} \int_0^L [f_{k+l}(x(s)) - f_{k+l}(x^*)]\, d(x' - x(s))^l, \tag{3}$$

where s denotes the length of that part of C between x^* and the variable point $x(s)$. If we carry out the differentiation indicated, we may write this as a sum of contour integrals

$$R_k(x'; x^*) = \sum_{\sigma_l = m - \sigma_k} \sum_{\sigma_j = 1} \int_{x^*}^{x'} [f_{k+l}(x) - f_{k+l}(x^*)] \frac{(x' - x)^{l-j}}{(l - j)!} (dx)^j. \tag{4}$$

In the case $n = 2$, $k = 0$, this equation, written out in full, is

$$\begin{aligned} R_{00}(x', y'; x^*, y^*) = \sum_{i+j=m} \int_{(x^*, y^*)}^{(x', y')} & [f_{ij}(x, y) - f_{ij}(x^*, y^*)] \\ & \cdot \left[\frac{(x' - x)^{i-1} (y' - y)^j}{(i-1)!\, j!}\, dx + \frac{(x' - x)^i (y' - y)^{j-1}}{i!\, (j-1)!}\, dy \right]. \end{aligned} \tag{5}$$

[4a] An example shows that no such formula holds for all curves.

To prove (3), let $x^0 = x^*, x^1, \cdots, x^p = x'$ be the end points of a subdivision of C. If in equation (6.3) of AE we subtract $f_k(x'')$ from both sides and change x, x', x'' to x^{i-1}, x^i, x' respectively, we find

$$R_k(x'; x^{i-1}) - R_k(x'; x^i) = \sum_{\sigma_l \leqq m-\sigma_k} \frac{R_{k+l}(x^i; x^{i-1})}{l!} (x' - x^i)^l. \tag{6}$$

As $R_k(x'_h; x^i_h) = 0$, summing over i gives

$$R_k(x'; x^*) = \sum_{\sigma_l \leqq m-\sigma_k} \frac{1}{l!} \sum_{i=1}^{p} R_{k+l}(x^i; x^{i-1})(x' - x^i)^l. \tag{7}$$

We show first that as the norm of the subdivision tends to zero, the terms on the right with $\sigma_l < m - \sigma_k$ tend to zero. By (2), given an $\eta > 0$, we can take the norm so small that

$$| R_j(x^i; x^{i-1}) | < r_{x^{i-1}x^i}\, \eta \qquad (\sigma_j < m).$$

As $| x'_h - x^i_h | \leqq L$, we find for $\sigma_l < m - \sigma_k$

$$\left| \sum_{i=1}^{p} R_{k+l}(x^i; x^{i-1})(x' - x^i)^l \right| < \eta L^{\sigma_l} \sum_{i=1}^{p} r_{x^{i-1}x^i} \leqq \eta L^{\sigma_l+1},$$

as η is arbitrary, the statement is proved.

Now take any l with $\sigma_l = m - \sigma_k$. We have

$$\sum_{i=1}^{p} R_{k+l}(x^i; x^{i-1})(x' - x^i)^l = \sum_{i=1}^{p} [f_{k+l}(x^i) - f_{k+l}(x^{i-1})]\,(x' - x^i)^l$$

$$= -\sum_{i=1}^{p-1} [f_{k+l}(x^i) - f_{k+l}(x^*)]\,[(x' - x^{i+1})^l - (x' - x^i)^l].$$

Putting this in (7) and passing to the limit as the norm of the subdivision tends to zero, we obtain (3).

LEMMA 3. *Let $f(x)$ be of class C^m on the curve C of length L with end points x^* and x'. If*

$$| f_k(x) - f_k(x^*) | < \epsilon \qquad (\sigma_k = m, x \text{ on } C), \tag{8}$$

then

$$| R_k(x'; x^*) | < n(m+1)^n L^{m-\sigma_k} \epsilon^{5} \qquad (\sigma_k \leqq m). \tag{9}$$

---

[5] The numerical factor $n(m+1)^n$ may of course be replaced by the factor

$$\sum_{\sigma_l = m-\sigma_k} \sum_{\sigma_j = 1} \frac{1}{(l-j)!} = \frac{n^{m-\sigma_k}}{(m-\sigma_k-1)!}.$$

This follows at once from (4) when we note that the first sum contains at most $(m+1)^n$ terms, the second sum contains n terms, and

$$\left| \int_{x^*}^{x'} \frac{(x'-x)^{l-j}}{(l-j)!} (dx)^j \right| < \int_0^L L^{\sigma_l - 1}\, ds = L^{\sigma_l}.$$

4. **Functions of class C^m in $R+B$.** One more lemma will lead us to the main theorem of the paper.

LEMMA 4. *If R has the property P and $f_l(x)$ is uniformly continuous in $R(\sigma_l = m)$, then $f_k(x)$ is also $(\sigma_k < m)$.*

Assuming this is true for values of k such that $\sigma_k > s$ $(0 \leqq s < m)$, we shall prove it for any k with $\sigma_k = s$. By hypothesis, $f_{k+l}(x)$ is uniformly continuous for $\sigma_l > 0$; from (1), we see that it is sufficient to show that for every $\epsilon > 0$ there is a $\delta > 0$ such that if x' and x^* are any two points of R with $r_{x'x^*} < \delta$, then $| R_k(x'; x^*) | < \epsilon$. Take $\delta' < 1$ so that $| f_k(x) - f_k(x^*) | < \epsilon/[n(m+1)^n]$ if $r_{xx^*} < \delta'$ and $\sigma_k = m$. Set $\delta = \delta'/\omega$, and take any two points x', x^* of R with $r_{x'x^*} < \delta$. There is a curve C joining them of length $L \leqq \omega r_{x'x^*} < \delta'$; if x is on C, then $r_{xx^*} < \delta'$. As $L < 1$, Lemma 3 gives $| R_k(x'; x^*) | < \epsilon$, as required.

THEOREM. *Let the region R have the property P, and let $f(x_1, \cdots, x_n)$ be of class C^m in R. If $\partial^{k_1+\cdots+k_n} f/\partial x_1^{k_1} \cdots \partial x_n^{k_n}$ $(k_1 + \cdots + k_n = m)$ can be defined on the boundary B of R so that it is continuous in $R+B$, then the definition of f can be extended throughout space so that it is of class C^m there.*

By the last lemma, $f_k(x)$ can be defined in B so that it is continuous in $R+B$ $(\sigma_k \leqq m)$. We must show that $f_0(x) = f(x)$ is of class C^m in $R+B$; the theorem then follows from Lemma 1. Given an $\epsilon > 0$, take $\delta' < 1$ so that $| f_k(x) - f_k(x^*) | < \epsilon/[n(m+1)^n \omega'^m]$ $(\sigma_k = m)$ if x and x^* are in $R+B$ and $r_{xx^*} < \delta'$; set $\delta = \delta'/\omega'$. Now let y' and y^* be any two points of $R+B$ with $r_{y'y^*} < \delta$. By Lemma 2, there is a curve C' joining them of length $L' \leqq \omega' r_{y'y^*} < \delta'$; all of C' except possibly its end points lies in R. If x' and x^* are interior points of C', and C, of length L, is that part of C' joining them, then, as $L < L'$, Lemma 3 gives

$$\frac{| R_k(x'; x^*) |}{r_{y'y^*}^{m-\sigma_k}} < n(m+1)^n \frac{L^{m-\sigma_k}}{r_{y'y^*}^{m-\sigma_k}} \frac{\epsilon}{n(m+1)^n \omega'^m} < \epsilon.$$

As $f_k(x)$ is continuous in $R+B$, (1) shows that $R_k(x'; x^*)$ is also. Hence, letting $x' \to y'$ and $x^* \to y^*$ in the above inequality, we obtain (2) for $R_k(y'; y^*)$, as required.

HARVARD UNIVERSITY.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

HARMONICS ON HOMOGENEOUS MANIFOLDS

BY H. WEYL

(Received April 9, 1934)

Introduction

Our investigation starts with the following data: a compact Lie group σ, a point field Π and a realization of σ by means of transformations (one-to-one correspondences) in Π, so that the element s of σ carries over the point P of Π into a point sP. $P, Q, \cdots$ denote arbitrary points of the point field, $s, t, \cdots$ arbitrary elements of the group. We assume that the group of transformations σ is transitive on Π so that all points P are equivalent with respect to σ. A function $f(P)$ is carried over by a transformation s of our group into the function sf defined by

$$sf(sP) = f(P) \qquad \text{or} \qquad sf(P) = f(s^{-1}P).$$

An h-dimensional *linear set* $\mathfrak{R}_h$ *of functions* $x_1\varphi_1(P) + \cdots + x_h\varphi_h(P)$ on Π with the basis consisting of the linearly independent functions $\varphi_1, \cdots, \varphi_h$ is *invariant* with respect to σ, if the transformed functions $s\varphi_\mu$ are linear combinations of the φ_μ:

$$s\varphi_\mu(P) = \sum_{\nu=1}^{h} u_{\nu\mu}(s)\cdot\varphi_\nu(P). \tag{1}$$

$U(s) = \| u_{\mu\nu}(s) \|$ then is a representation U of the group σ; we say that the invariant set *belongs to*, or *is associated with, the representation* U. The set is *primitive* if the corresponding representation is irreducible.

The classical example of the situation at hand is the sphere Π under the influence of the group σ of all rotations. Here the *spherical harmonics* of a given order form a primitive invariant set. In accordance with this example one will expect that *the primitive invariant sets form together a complete orthogonal system on* Π. We may stress the analogy by using the term *harmonic set of order* U instead of invariant set associated with the representation U.

Every abstract group σ allows a realization ("regular realization") the point field of which is the group manifold σ itself, and which represents the group element a by the transformation $s \to as$. In this case the columns of the irreducible representations are the primitive harmonic sets: thus an irreducible representation of degree h yields h linearly independent harmonic sets associated with it. For this case orthogonality was proved by Schur, completeness by Peter and the author.[1] The method can be carried over, practically

[1] I. Schur, Sitzungsber. Berl. Akad. 1905, p. 406; 1924, p. 199. Peter and Weyl, Math. Ann. **97** (1927), p. 737 (cited as PW).

without any change, to our more general problem which was first propounded and treated by E. Cartan.[2] It becomes in this case even more lucid since the coincidence of the point field with the group manifold in the special case is embarrassing rather than simplifying. Two observations caused me to come back to this problem after Cartan's paper:

1) It is not necessary to assume Π to bear a metric, for by means of the invariant volume measure on the group manifold one is able to form the mean value of a continuous function $f(P)$ on Π in the following way:

$$m(P) = \int_s f(s^{-1} P).$$

$\int_s$ denote the average with respect to s. We have

$$m(a^{-1} P) = \int_s f(s^{-1} a^{-1} P) = \int_t f(t^{-1} P) = m(P) \qquad [t = as].$$

Since every point Q can be derived from P by means of a transformation a^{-1} of the group, $m(Q)$ is equal to $m(P)$ for any two points P and Q; hence $m(P)$ is a constant m, and this constant may be called the *mean value* of $f(P)$:

$$m = \int_P f(P).$$

2) Cartan proves completeness in an indirect way, making use of kernels $F(P, Q)$ which are invariant for the transformations of the group:

$$F(sP, sQ) = F(P, Q). \tag{2}$$

But one gets the relation of completeness holding for the function $f(P)$ quite directly if one, in even closer accordance with the procedure of Peter and myself, chooses

$$F(P, Q) = \int_s \bar{f}(s^{-1} P)\, f(s^{-1} Q) \tag{3}$$

or, what amounts to the same thing, one considers, following E. Schmidt's theory of the eigen-values of asymmetric kernels, the two Hermitian conjugate operations

$$\left.\begin{aligned} f\ \varphi(P) \to \psi(s) &= \int_P f(s, P)\, \varphi(P) \\ f^*\ \psi(s) \to \varphi(P) &= \int_s \bar{f}(s, P)\, \psi(s) \end{aligned}\right\} \quad \text{with } f(s, P) = f(s^{-1} P). \tag{4}$$

[2] Rend. Circ. Mat. di Palermo **53** (1929), p. 217.

(The notation of course is not intended to suggest that the two operators are reciprocal.)

Indeed, if $\gamma \neq 0$ is an h-fold eigen-value of $F(P, Q)$ with the h independent eigen-functions $\bar{\varphi}_\mu(P)$, chosen as a normalized, unitary-orthogonal system:

$$\int_P \bar{\varphi}_\mu(P)\, \varphi_\nu(P) = \delta_{\mu\nu},$$

then, on account of the invariance of F, $\bar{\varphi}_\mu(sP)$ is an eigen-function too, and $\{\varphi_1(P), \cdots, \varphi_h(P)\}$ is an invariant set of functions, eq. (1), associated with a definite representation U. The φ being normalized, this U is unitary.

$$\alpha_\mu = \int_P f(P)\, \bar{\varphi}_\mu(P) \tag{5}$$

may denote the corresponding Fourier coefficients of f. A simple computation which is carried through in §2 yields

$$h\gamma = \alpha_1 \bar{\alpha}_1 + \cdots + \alpha_h \bar{\alpha}_h. \tag{6}$$

Consequently, the theorem that the sum of the eigen-values of the kernel $F(P, Q)$ equals its trace, $\mathrm{tr}\,(F) = \int_P F(P, P)$, leads immediately to the relation of completeness, for we have

$$F(P, P) = \int_s |\, f(sP)\,|^2 = \int_P |\, f(P)\,|^2, \qquad \mathrm{tr}\,(F) = \int_P |\, f(P)\,|^2.$$

These preliminary remarks were intended for the expert who is familiar with Cartan's paper. I now embark upon the systematic exposition by which I should like to replace the two papers Peter-Weyl and Cartan. Concerning the group, we make those assumptions which grant the existence of a differential invariant volume measure on the group manifold (cf. PW). $P' = sP$ is to depend uniformly continuously on s and P. In the first part we discuss *orthogonality* in the same manner as Cartan discussed it; it is shown that at most h independent harmonic sets can be associated with an irreducible representation of degree h. The second part contains the proof of *completeness* and the *expansion* and *approximation theorems*.

I may point out, by the way, that the whole investigation could be carried through within the more general frame of *von Neumann's theory of almost periodic functions on arbitrary groups*.[3] Here group and point field need not be compact, nor even topological. On the contrary, we topologize the group manifold and the point field relatively to the given function f by agreeing to say that two group elements s and t, or two points P and Q, have a distance $\leqq \epsilon$ if in the first case $|\, f(sP) - f(tP)\,| \leqq \epsilon$ identically with respect to P, and if in the second case $|\, f(sP) - f(sQ)\,| \leqq \epsilon$ identically with respect to s. Following

[3] Which will soon appear in the Trans. of the Amer. Math. Soc.

Bochner and von Neumann, f is called *almost periodic* if, in consequence of this topologizing, the two manifolds σ and Π become "*finite*" in the sense that, to every positive ϵ, one can choose a finite number of points such that every point has a distance $\leqq \epsilon$ from one of these chosen points. (A "finite" country can be watched by a finite number of policemen, however small the radius of action of the single policeman may be!) Following von Neumann, $sf(P) = f(s^{-1}P)$, considered as a function of s, then has a mean value; this is independent of P and may be taken as the mean value of the almost periodic function $f(P)$ on Π. When looked upon from this more general viewpoint, the present note should absorb my paper on integral equations and almost periodic functions, Math. Ann. **97** (1926), p. 338, as well as von Neumann's forthcoming paper on almost periodic functions on groups.

§1. Orthogonality

Let $\varphi(P)$ run through a harmonic set $\mathfrak{R}_h$ of continuous functions

$$x_1 \varphi_1(P) + \cdots + x_h \varphi_h(P) \tag{7}$$

with the unitary basis $\varphi_\mu(P)$:

$$\int \bar{\varphi}_\mu(P)\, \varphi_\nu(P) = \delta_{\mu\nu}.$$

According to (1) it belongs to a definite unitary representation U: the function (7) with the parameters x_μ is carried over into the function with the parameters

$$x'_\mu = \sum_\nu u_{\mu\nu}(s)\, x_\nu$$

by means of the transformation s. If one makes a unitary change of basis, the representation is changed into a unitary-equivalent one. The harmonic set is called *primitive* provided the corresponding representation is irreducible. For well-known reasons of elementary unitary geometry it follows that a unitary representation can be completely decomposed into irreducible constituents; by means of the corresponding choice of basis $\mathfrak{R}_h$ breaks up into a number of mutually orthogonal *primitive* harmonic sets.

THEOREM 1. *The functions of two primitive sets belonging to inequivalent representations are mutually orthogonal.*

Let the two sets belonging to the inequivalent irreducible representations U and V consist of the functions

$$\varphi(P) = x_1 \varphi_1(P) + \cdots + x_h \varphi_h(P) \qquad \text{and}$$

$$\varphi'(P) = y_1 \varphi'_1(P) + \cdots + y_k \varphi'_k(P) \text{ respectively.}$$

$$\int_P \bar{\varphi}(P)\, \varphi'(P) = \sum_{\mu,\,\nu} a_{\mu\nu}\, \bar{x}_\mu\, y_\nu \tag{8}$$

is invariant with respect to the substitutions $U(s)$ and $V(s)$ simultaneously performed on x and y:

$$U^*(s)AV(s) = A \qquad \text{or} \qquad AV(s) = U(s)A$$

A is the matrix of the $a_{\mu\nu}$, A^* in general the transposed-conjugate of A. In accordance with I. Schur's well-known lemma, this yields:

$$A = 0, \qquad a_{\mu\nu} = 0 \qquad \text{or} \qquad \int \bar{\varphi}_\mu \varphi'_\nu = 0.$$

THEOREM 2. *Two primitive harmonic sets $\varphi_\mu(P)$ and $\varphi'_\mu(P)$ associated with the same representation U satisfy the relations*

$$\int_P \bar{\varphi}_\mu(P)\, \varphi'_\nu(P) = c\delta_{\mu\nu}.$$

The constant c is determined by

$$(\varphi, \varphi') \equiv \frac{1}{h} \int_P \sum_{\mu=1}^{h} \bar{\varphi}_\mu(P)\, \varphi'_\mu(P) = c. \tag{9}$$

Again one forms (8). Its coefficient matrix A commutes with all $U(s)$ and is therefore, according to Schur's lemma, a multiple c of the unit matrix.[4]

Consequences of Theorem 2: Being given several primitive harmonic sets $\varphi^{(1)}, \varphi^{(2)}, \cdots$ associated with the same representation U, every linear combination $a^{(1)}\varphi^{(1)} + a^{(2)}\varphi^{(2)} + \cdots$ is such a set again. Within this linear manifold one may define a Hermitian metric by means of the scalar product (9), and hence one may choose the linearly independent among the sets such that $(\varphi^{(i)}, \varphi^{(k)}) = \delta_{ik}$. Theorem 2 then shows that *all* the functions of *all* the sets are orthogonal to each other:

$$\int_P \bar{\varphi}_\mu^{(i)}(P)\, \varphi_\nu^{(k)}(P) = \delta_{ik}\delta_{\mu\nu} = \begin{cases} 1 \text{ for } i = k,\ \mu = \nu \\ 0 \text{ in all other cases.} \end{cases}$$

This implies the statement: being given several linearly independent harmonic sets associated with the same irreducible representation U — so that no non-trivial linear relation

$$\sum_i a^{(i)} \varphi_\mu^{(i)}(P) = 0$$

with constant coefficients $a^{(i)}$ exists — all the functions of the sets are linearly independent; that is, there does not exist any non-trivial relation of the form

[4] Theorem 2 implies the fact: An invariant set $\{\varphi_\mu(P)\}$ belonging to the *unitary irreducible* representation U satisfies equations

$$\int \bar{\varphi}_\mu(P)\, \varphi_\nu(P) = c\delta_{\mu\nu}.$$

Therefore only two cases are possible: either all the $\varphi_\mu(P)$ vanish ($c = 0$) or the $\varphi_\mu(P)$ are linearly independent and orthogonal (with a "normalizing" constant $c \neq 0$ however, instead of $c = 1$).

$$\sum_{i,\mu} a_\mu^{(i)} \varphi_\mu^{(i)}(P) = 0.$$

THEOREM 3. *There exist at most h linearly independent primitive harmonic sets associated with a given representation of degree h.*

Consider the transformed functions of the given sets:

$$s\varphi_\mu^{(i)}(P) = \varphi_\mu^{(i)}(s^{-1}P)$$

as functions of s. These functions of s are linearly independent as well as the functions $\varphi_\mu^{(i)}(P)$ of P, since P can be carried over into every point by means of a transformation s^{-1}. But these functions of s are, according to the equations

$$s\varphi_\mu^{(i)}(P) = \sum_\nu u_{\nu\mu}(s)\, \varphi_\nu^{(i)}(P),$$

linear combinations of the h^2 functions $u_{\nu\mu}(s)$. Hence their number cannot be greater than h^2 and the number of the sets cannot be greater than h.

The harmonic sets associated with a given irreducible representation of degree h form the elements of a linear family of $l \leqq h$ dimensions. The basis $\varphi^{(i)}$ $(i = 1, \cdots, l)$ within this family may always be chosen unitary according to the metric (9):

$$(\varphi^{(i)}, \varphi^{(k)}) = \delta_{ik}.$$

By this requirement the basis is determined but for an l-dimensional unitary transformation: a definite choice of the basis is called a "*normalization.*"

Application to the regular realization: Let U be an irreducible representation of σ of degree h. The equation

$$U(a^{-1}s) = U(a^{-1})\, U(s) = U^{-1}(a)\, U(s) = U^*(a)\, U(s)$$

states that the columns $\varphi^{(1)}(s), \cdots, \varphi^{(h)}(s)$ of $U(s)$: $\varphi_\mu^{(i)}(s) = u_{\mu i}(s)$ are invariant sets on the group manifold associated with the representation $\overline{U}$; so in this case the maximum possible number h of existing linear invariant sets is reached. Since we have

$$\sum_\nu \bar{\varphi}_\nu^{(i)}(s)\, \varphi_\nu^{(k)}(s) = \delta_{ik}$$

($U(s)$ being unitary) there follows, in using the notation (9):

$$(\varphi^{(i)}, \varphi^{(k)}) = \frac{1}{h}\, \delta_{ik},$$

and hence the relations of orthogonality

$$\int_s \bar{u}_{\mu i}(s)\, u_{\nu k}(s) = \frac{1}{h}\, \delta_{\mu\nu}\, \delta_{ik},$$

first derived by I. Schur, obtain. In accordance with Theorem 1, the elements $u_{\mu i}(s)$, $v_{\nu k}(s)$ of the matrices $U(s)$, $V(s)$ of two inequivalent irreducible representations are unitary-orthogonal to each other.

is invariant with respect to the substitutions $U(s)$ and $V(s)$ simultaneously performed on x and y:

$$U^*(s)AV(s) = A \qquad \text{or} \qquad AV(s) = U(s)A$$

A is the matrix of the $a_{\mu\nu}$, A^* in general the transposed-conjugate of A. In accordance with I. Schur's well-known lemma, this yields:

$$A = 0, \qquad a_{\mu\nu} = 0 \qquad \text{or} \qquad \int \bar{\varphi}_\mu \varphi'_\nu = 0.$$

THEOREM 2. *Two primitive harmonic sets $\varphi_\mu(P)$ and $\varphi'_\mu(P)$ associated with the same representation U satisfy the relations*

$$\int_P \bar{\varphi}_\mu(P)\, \varphi'_\nu(P) = c\delta_{\mu\nu}.$$

The constant c is determined by

$$(\varphi, \varphi') \equiv \frac{1}{h} \int_P \sum_{\mu=1}^{h} \bar{\varphi}_\mu(P)\, \varphi'_\mu(P) = c. \tag{9}$$

Again one forms (8). Its coefficient matrix A commutes with all $U(s)$ and is therefore, according to Schur's lemma, a multiple c of the unit matrix.[4]

Consequences of Theorem 2: Being given several primitive harmonic sets $\varphi^{(1)}, \varphi^{(2)}, \cdots$ associated with the same representation U, every linear combination $a^{(1)}\varphi^{(1)} + a^{(2)}\varphi^{(2)} + \cdots$ is such a set again. Within this linear manifold one may define a Hermitian metric by means of the scalar product (9), and hence one may choose the linearly independent among the sets such that $(\varphi^{(i)}, \varphi^{(k)}) = \delta_{ik}$. Theorem 2 then shows that *all* the functions of *all* the sets are orthogonal to each other:

$$\int_P \bar{\varphi}_\mu^{(i)}(P)\, \varphi_\nu^{(k)}(P) = \delta_{ik}\delta_{\mu\nu} = \begin{cases} 1 \text{ for } i = k,\ \mu = \nu \\ 0 \text{ in all other cases.} \end{cases}$$

This implies the statement: being given several linearly independent harmonic sets associated with the same irreducible representation U — so that no non-trivial linear relation

$$\sum_i a^{(i)} \varphi_\mu^{(i)}(P) = 0$$

with constant coefficients $a^{(i)}$ exists — all the functions of the sets are linearly independent; that is, there does not exist any non-trivial relation of the form

[4] Theorem 2 implies the fact: An invariant set $\{\varphi_\mu(P)\}$ belonging to the *unitary irreducible* representation U satisfies equations

$$\int \bar{\varphi}_\mu(P)\, \varphi_\nu(P) = c\delta_{\mu\nu}.$$

Therefore only two cases are possible: either all the $\varphi_\mu(P)$ vanish ($c = 0$) or the $\varphi_\mu(P)$ are linearly independent and orthogonal (with a "normalizing" constant $c \neq 0$ however, instead of $c = 1$).

$$\sum_{i,\mu} a_\mu^{(i)} \varphi_\mu^{(i)}(P) = 0.$$

THEOREM 3. *There exist at most h linearly independent primitive harmonic sets associated with a given representation of degree h.*

Consider the transformed functions of the given sets:

$$s\varphi_\mu^{(i)}(P) = \varphi_\mu^{(i)}(s^{-1}P)$$

as functions of s. These functions of s are linearly independent as well as the functions $\varphi_\mu^{(i)}(P)$ of P, since P can be carried over into every point by means of a transformation s^{-1}. But these functions of s are, according to the equations

$$s\varphi_\mu^{(i)}(P) = \sum_\nu u_{\nu\mu}(s)\, \varphi_\nu^{(i)}(P),$$

linear combinations of the h^2 functions $u_{\nu\mu}(s)$. Hence their number cannot be greater than h^2 and the number of the sets cannot be greater than h.

The harmonic sets associated with a given irreducible representation of degree h form the elements of a linear family of $l \leqq h$ dimensions. The basis $\varphi^{(i)}$ $(i = 1, \cdots, l)$ within this family may always be chosen unitary according to the metric (9):

$$(\varphi^{(i)}, \varphi^{(k)}) = \delta_{ik}.$$

By this requirement the basis is determined but for an l-dimensional unitary transformation: a definite choice of the basis is called a "*normalization.*"

Application to the regular realization: Let U be an irreducible representation of σ of degree h. The equation

$$U(a^{-1}s) = U(a^{-1})\, U(s) = U^{-1}(a)\, U(s) = U^*(a)\, U(s)$$

states that the columns $\varphi^{(1)}(s), \cdots, \varphi^{(h)}(s)$ of $U(s)$: $\varphi_\mu^{(i)}(s) = u_{\mu i}(s)$ are invariant sets on the group manifold associated with the representation $\overline{U}$; so in this case the maximum possible number h of existing linear invariant sets is reached. Since we have

$$\sum_\nu \bar{\varphi}_\nu^{(i)}(s)\, \varphi_\nu^{(k)}(s) = \delta_{ik}$$

($U(s)$ being unitary) there follows, in using the notation (9):

$$(\varphi^{(i)}, \varphi^{(k)}) = \frac{1}{h}\, \delta_{ik},$$

and hence the relations of orthogonality

$$\int_s \bar{u}_{\mu i}(s)\, u_{\nu k}(s) = \frac{1}{h}\, \delta_{\mu\nu}\, \delta_{ik},$$

first derived by I. Schur, obtain. In accordance with Theorem 1, the elements $u_{\mu i}(s)$, $v_{\nu k}(s)$ of the matrices $U(s)$, $V(s)$ of two inequivalent irreducible representations are unitary-orthogonal to each other.

§2. Completeness

Let $f(P)$ be a continuous function on II. We call the mean value $\int_P \bar{f}(P) f(P)$ the *magnitude* $||f||^2$ of f. We form its *Fourier coefficients*

$$\int_P f(P)\, \bar{\varphi}_\mu(P) = \alpha_\mu$$

with respect to a given primitive harmonic set

$$\mathfrak{R}_h: \qquad \varphi_1(P), \cdots, \varphi_h(P)$$

of orthogonal functions.

$$e(P) = \alpha_1 \varphi_1(P) + \cdots + \alpha_h \varphi_h(P) \tag{10}$$

is the orthogonal projection of f upon $\mathfrak{R}_h$; for we have

$$\int_P (f(P) - e(P))\, \bar{\varphi}_\mu(P) = 0. \tag{11}$$

$$\alpha_1 \bar{\alpha}_1 + \cdots + \alpha_h \bar{\alpha}_h, \tag{12}$$

the magnitude of e, is the "*contribution*" of $\mathfrak{R}_h$ to the magnitude of f. If one considers several independent primitive harmonic sets (*1*), (*2*), (*3*), $\cdots$ —which may be chosen in orthogonal normalization as far as they belong to the same representation—the sum of the individual contributions (*1*), (*2*), (*3*), $\cdots$ will certainly not surpass the total magnitude of f, according to Bessel's inequality. The theorem of completeness maintains that here the sign of equality holds, provided the sum be extended over a complete system of normalized primitive sets. This implies the fact that there exist sets occurring in f, if f is not identically zero. A set $\mathfrak{R}_h$ "occurs" in f, if its contribution (12) is different from zero; we shall say that $\mathfrak{R}_h$ occurs in f with a "*weight*"

$$= \frac{1}{h} (\alpha_1 \bar{\alpha}_1 + \cdots + \alpha_h \bar{\alpha}_h).$$

One proves the theorem of completeness by constructing the eigen-values and eigen-functions of the couple of Hermitian conjugate operators f, f^*, formula (4), following E. Schmidt. We are dealing with an eigen-function $\varphi(P)$ corresponding to the eigen-value γ, if $\varphi(P) \to \psi(s) \to \gamma \cdot \varphi(P)$, by performing first f and then f^*. At the same time $\psi(s)$ is carried over into $\gamma \cdot \psi(s)$ by the successive operations f^*, f (in reversed order). It is not quite, but nearly, true that every primitive harmonic set consists of eigen-functions of the couple (f, f^*)—so that the eigen-functions of such a couple are "essentially" independent of the nature of the function $f(P)$. The statement becomes correct if one normalizes the basis of harmonic sets associated with the same irreducible representation in an appropriate manner *dependent on f*. The proof rests upon the following fact.

The harmonic set $\varphi_\mu(P)$ of order U gives rise to a harmonic set $\hat{\varphi}_\mu(P)$ of the same order U by applying the operators f, f^*:

$$\varphi_\mu(P) \to \psi_\mu(s) \to \hat{\varphi}_\mu(P).$$

Indeed, we have

$$\hat{\varphi}_\mu(P) = \int_Q F(P, Q)\, \varphi_\mu(Q)$$

where $F = f^*f$ arises from f and f^* by means of "convolution":

$$F(P, Q) = \int_s f^*(P, s)\, f(s, Q) = \int_s \bar{f}(s^{-1} P)\, f(s^{-1} Q) \qquad \text{[compare (3)]}.$$

Hence

$$\hat{\varphi}_\mu(aP) = \int_Q F(aP, Q)\, \varphi_\mu(Q) = \int_Q F(aP, aQ)\, \varphi_\mu(aQ) = \int_Q F(P, Q)\, \varphi_\mu(aQ).$$

Consequently the $\hat{\varphi}_\mu(aP)$ are connected with the functions $\hat{\varphi}_\mu(P)$ by means of the same linear transformation $U(a^{-1})$ by which $\varphi_\mu(P)$ is carried over into $\varphi_\mu(aP)$.

Hence, $\varphi^{(i)}(P)$ $[i = 1, \cdots, l]$ being the l harmonic sets of the primitive order U in normalized form, equations like

$$\hat{\varphi}^{(i)}(P) = \sum_{k=1}^{l} c_{ik}\, \varphi^{(k)}(P)$$

will hold with constant coefficients c_{ik}:

$$c_{ik} = (\varphi^{(k)}, \hat{\varphi}^{(i)}) = \frac{1}{h} \sum_\mu \int_Q \int_P \bar{\varphi}_\mu^{(k)}(P)\, F(P, Q)\, \varphi_\mu^{(i)}(Q).$$

$\| c_{ik} \|$ is a Hermitian matrix: $\bar{c}_{ki} = c_{ik}$, since $F(P, Q)$ is a Hermitian kernel: $\bar{F}(Q, P) = F(P, Q)$. By appropriate choice of the unitary basis $\| c_{ik} \|$ may be reduced to the form

$$\begin{Vmatrix} \gamma_1 & & & \\ & \gamma_2 & & \\ & & \ddots & \\ & & & \gamma_l \end{Vmatrix};$$

under such circumstances the set $\varphi^{(i)}(P)$ consists of eigen-functions of the couple (f, f^*) corresponding to the eigen-value γ_i.

For this reason we look for the primitive harmonic sets occurring in f by constructing, according to E. Schmidt, the eigen-values of this kernel arranged in decreasing order, and the corresponding eigen-functions. The essential steps may be briefly repeated as follows. We suppose that f is not identically zero. One forms the trace Γ^n of F^n, that is of the n-fold iterated kernel F. In particular

$$\Gamma^1 = \int_P |f(P)|^2 = \|f\|^2, \qquad \Gamma^2 = \int_P \int_Q |F(P, Q)|^2.$$

Γ^2 is > 0 since $F(P, P) = \Gamma^1 > 0$. We have the inequalities[5]

$$\Gamma^{m+n} \leqq \Gamma^m \cdot \Gamma^n, \qquad (\Gamma^n)^2 \leqq \Gamma^{n-1} \cdot \Gamma^{n+1}. \tag{13}$$

Hence Γ^{n+1}/Γ^n increases monotonely towards a positive limit γ which is less or equal to $\sqrt[m]{\Gamma^m}$, in particular $\leqq \Gamma^1 = \|f\|^2$. For later use we note the inequality

$$\Gamma^2 \leqq \gamma \cdot \Gamma^1. \tag{14}$$

$\Gamma^n/\gamma^n \geqq 1$ decreases monotonely towards a limit $h \geqq 1$, and $F^n(P, Q)/\gamma^n$ tends uniformly towards a limit $E(P, Q)$. Indeed, in writing $F^{n+1} = f^* G^n f$ one gets from Schwarz's inequality

$$\left| \frac{F^{n+1}}{\gamma^{n+1}} - \frac{F^{m+1}}{\gamma^{m+1}} \right|^2 \leqq \frac{1}{\gamma^2} \cdot \int_s \int_t \left| \frac{G^n(s, t)}{\gamma^n} - \frac{G^m(s, t)}{\gamma^m} \right|^2 = \frac{1}{\gamma^2} \left(\frac{\Gamma^{2n}}{\gamma^{2n}} - 2 \frac{\Gamma^{n+m}}{\gamma^{n+m}} + \frac{\Gamma^{2m}}{\gamma^{2m}} \right).$$

(Here, for convenience' sake, we supposed $\|f\|^2$ to be $\leqq 1$.) $E(P, Q)$ is an invariant Hermitian kernel for which the following relations hold:

$$FE = \gamma E, \qquad EE^* = EE = E, \qquad \operatorname{tr}(E) = h \geqq 1. \tag{15}$$

On account of the second and third equations, one gets an expression

$$E(P, Q) = \bar{\varphi}_1(P)\, \varphi_1(Q) + \cdots + \bar{\varphi}_h(P)\, \varphi_h(Q) \tag{16}$$

where the $\bar{\varphi}$ form the unitary-orthogonal system of the h eigen-functions of F corresponding to the eigen-value γ (consequently $h \geqq 1$ ultimately proves to be an integer!). $\varphi_\mu(s^{-1} P)$ is necessarily a linear combination of the $\varphi_\mu(P)$ themselves: the $\varphi_\mu(P)$ therefore form a harmonic set corresponding to a definite representation U of degree h. For, one has

$$\varphi_\mu(Q) = \int_P \varphi_\mu(P)\, E(P, Q),$$

hence

$$\varphi_\mu(s^{-1} Q) = \int_P \varphi_\mu(P)\, E(P, s^{-1} Q) = \int_P \varphi_\mu(s^{-1} P)\, E(s^{-1} P, s^{-1} Q)$$
$$= \int_P \varphi_\mu(s^{-1} P)\, E(P, Q) = \sum_\nu \varphi_\nu(Q) \int_P \varphi_\mu(s^{-1} P)\, \bar{\varphi}_\nu(P).$$

---

[5] For the proof of the following inequalities cf. Weyl, Math. Ann. **97** (1926), pp. 342–343. It is immaterial that the arguments of some of the kernels here considered vary within different fields (σ and Π). Of course, a trace like $\operatorname{tr}(F) = \int_P F(P, P)$ can be formed only from a kernal $F(P, Q)$ the arguments of which vary in the *same* field. In proving the inequalities (13) one has to make use of the fact that the trace $\operatorname{tr}(f^* f \cdots f^* f)$ of the $2n$-term convolution is not changed when one shifts the factors cyclically by one place.

The set $\{\varphi_\mu(P)\}$—if it is not primitive itself—can be split into primitive mutually orthogonal harmonic sets.

In performing first the operation f, then f^* [compare eq. (4)], we have

$$\bar{\varphi}_\mu(P) \to \psi_\mu(s) \to \gamma\bar{\varphi}_\mu(P). \tag{17}$$

The $\psi_\mu(s)$ are orthogonal on σ; for,

$$\int_s \bar{\psi}_\mu(s)\,\psi_\nu(s) = \int_P\int_Q \varphi_\mu(P)\,F(P,Q)\,\bar{\varphi}_\nu(Q) = \gamma\int_P \varphi_\mu(P)\,\bar{\varphi}_\nu(P) = \gamma\delta_{\mu\nu}. \tag{18}$$

And just as the $\bar{\varphi}_\mu$ are eigen-functions of $F = f^*f$ corresponding to the eigen-value γ, so are the $\psi_\mu(s)$ eigen-functions of $G = ff^*$ corresponding to the same eigen-value. From (18) there follows

$$\int_s \sum_\mu \bar{\psi}_\mu(s)\,\psi_\mu(s) = h\gamma. \tag{19}$$

One finds

$$\psi_\mu(s) = \int_P f(s^{-1}P)\,\bar{\varphi}_\mu(P) = \int_P f(P)\,\bar{\varphi}_\mu(sP)$$

$$= \sum_\nu \bar{u}_{\nu\mu}(s^{-1})\int_P f(P)\,\bar{\varphi}_\nu(P) = \sum_\nu u_{\mu\nu}(s)\,\alpha_\nu.$$

Since $U(s)$ is unitary, one obtains from this expression:

$$\sum_\mu \bar{\psi}_\mu(s)\,\psi_\mu(s) = \alpha_1\bar{\alpha}_1 + \cdots + \alpha_h\bar{\alpha}_h. \tag{20}$$

Comparison of (19) and (20) yields the relation (6) which is decisive for our whole investigation. γ being > 0, *the invariant set* $\{\varphi_\mu(s)\}$ *actually occurs in* f, namely with the weight γ. The same holds for any of the partial primitive harmonic sets into which the whole set may be decomposed.

The next step in constructing all primitive harmonic sets occurring in f consists in deducting from f the contribution just found out:

$$f_1(P) = f(P) - e(P) = f(P) - \{\alpha_1\varphi_1(P) + \cdots + \alpha_h\varphi_h(P)\} \tag{21}$$

and in applying the same construction on f_1 which we carried through for f (provided f_1 does not vanish identically). In this way we get an invariant set $\varphi^{(1)}(P)$ consisting of $h_1 \geqq 1$ terms and occurring in f_1 with a weight $\gamma_1 > 0$; the h_1 conjugate functions $\bar{\varphi}^{(1)}$ are eigen-functions of the kernel $F_1 = f_1^*f_1$ corresponding to the eigen-value γ_1. It is important to realize: 1) that the new functions $\varphi^{(1)}$ are orthogonal to the previous φ. Hence they occur in f with the same weight γ_1 as in f_1 and the $\bar{\varphi}^{(1)}$ are eigen-functions of the kernel F corresponding to the eigen-value γ_1. 2) $\gamma_1 < \gamma$.

We have

$$\int_s\int_Q \bar{f}_1(s^{-1}P)\,f_1(s^{-1}Q)\,\bar{\varphi}_\nu^{(1)}(Q) = \gamma_1\cdot\bar{\varphi}_\nu^{(1)}(P).$$

In multiplying the left side by $\varphi_\mu(P)$ and averaging over P, we get zero; for

$$\int_P \bar{f}_1(s^{-1}P)\,\varphi_\mu(P) = \int_P \bar{f}_1(P)\,\varphi_\mu(sP) \text{ is } = 0,$$

since $\varphi_\mu(sP)$ is a linear combination of the $\varphi(P)$ and since the equation

$$\int_P \bar{f}_1(P)\,\varphi_\mu(P) = 0$$

holds for every $\varphi_\mu(P)$. Consequently, as we maintained before:

$$\gamma_1 \cdot \int_P \varphi_\mu(P)\,\bar{\varphi}_\nu^{(1)}(P) = 0.$$

The subtracted function $e(P)$ gives rise to a corresponding kernel

$$e(s^{-1}P) = e(s, P)$$

which we compute to be

$$(22) \qquad e(s,P) = \sum_\nu \alpha_\nu\,\varphi_\nu(s^{-1}P) = \sum_{\nu,\,\mu} \alpha_\nu\,u_{\mu\nu}(s)\,\varphi_\mu(P) = \sum_\mu \psi_\mu(s)\,\varphi_\mu(P).$$

From this we derive following (17): $f^*e = \gamma E$, and since we have, in addition to that, $e^*e = \gamma E$ following (18), we get $f_1^*\,e = 0$, $e^*f_1 = 0$, and hence $f^*f = f_1^*f_1 + \gamma E$ or $F = \gamma E + F_1$. The last equation shows in the simplest manner that the $\varphi^{(1)}$, being orthogonal to the φ's, are eigen-functions of F as well as of F_1. Furthermore, iteration will give

$$F^n = \gamma^n E + F_1^n, \qquad \Gamma^n = h \cdot \gamma^n + \Gamma_1^n.$$

Hence Γ_1^n/γ^n tends towards zero with increasing n, and consequently the number γ_1 for which Γ_1^n/γ_1^n tends towards the limit $h_1 \geqq 1$ is necessarily $< \gamma$.

After applying the construction p times, one is left with a certain remainder f_p. We have to prove—provided the construction does not come to an end by f_p becoming identically zero for a certain finite p—that

$$\Gamma_p^1 = \int_P |f_p(P)|^2 = \mathrm{tr}\,(f_p^* f_p)$$

tends towards zero with increasing p. From the relation (Bessel's inequality)

$$(23) \qquad (h + h_1 + \cdots + h_p)\,\gamma_p \leqq h\gamma + h_1\gamma_1 + \cdots + h_p\gamma_p$$
$$= \|e\|^2 + \|e_1\|^2 + \cdots + \|e_p\|^2 \leqq \|f\|^2 \leqq 1$$

we first conclude only that γ_p tends to zero with increasing p. The connection with the result wanted is established by means of the inequality (14):

$$(24) \qquad \Gamma_p^2 \leqq \gamma_p \cdot \Gamma_p^1.$$

Let $\mathfrak{V}_\epsilon$ denote the domain on the group manifold to which s belongs if

$$\int_P |f(s^{-1}P) - f(P)|^2 \leqq \epsilon^2. \tag{25}$$

One cannot be sure that $\mathfrak{V}_\epsilon$ has a volume in the Jordan sense, but we certainly can assign a positive lower bound V_ϵ for its volume. After division by the volume V of the whole group manifold, the fraction $w_\epsilon = V_\epsilon/V$ is a lower bound of the "probability" of s satisfying the inequality (25). I maintain: As soon as $\gamma_p \leqq \epsilon^2 \cdot w_\epsilon$,

$$||f_p||^2 = ||f||^2 - ||e||^2 - ||e_1||^2 - \cdots - ||e_{p-1}||^2 \leqq 4\epsilon^2.$$

PROOF:

$$G_p = f_p f_p^* = f f_p^*,$$

hence

$$G_p(s) \equiv G_p(s, 1) = \int_P f(s^{-1}P)\, \bar{f}_p(P).$$

(1 denotes the unit element on σ.)

$\Gamma_p^1 = \int_P |f_p(P)|^2 = G_p(1)$ may be put $= \delta_p^2$. Suppose it to be $> 4\epsilon^2$. Then

$$G_p(1) > 2\epsilon\delta_p. \tag{26}$$

On the other hand,

$$G_p(s) - G_p(1) = \int_P \{f(s^{-1}P) - f(P)\}\, \bar{f}_p(P),$$

and consequently, by means of Schwarz's inequality:

$$|G_p(s) - G_p(1)|^2 \leqq \int_P |f(s^{-1}P) - f(P)|^2 \cdot \int_P |f_p(P)|^2 \leqq \epsilon^2 \cdot \delta_p^2 \text{ in } \mathfrak{V}_\epsilon. \tag{27}$$

We then have in $\mathfrak{V}_\epsilon$, according to (26) and (27):

$$|G_p(s)| > \epsilon\delta_p,$$

and therefore

$$\Gamma_p^2 = \int_s |G_p(s)|^2 > \epsilon^2 \delta_p^2 \cdot w_\epsilon.$$

This is in contradiction with the inequality (24) or $\Gamma_p^2 \leqq \gamma_p \delta_p^2$ as soon as $\gamma_p \leqq \epsilon^2 w_\epsilon$.

Let us number the orthogonal functions in the order in which we got them, without marking the individual sets:

$$\varphi_1(P), \varphi_2(P), \cdots$$

and let us now cancel the supplementary condition $\|f\|^2 \leqq 1$ again. Then, in taking account of the inequality (23), our final result reads as follows:

$$(28) \qquad \|f\|^2 - (\alpha_1\bar{\alpha}_1 + \alpha_2\bar{\alpha}_2 + \cdots + \alpha_n\bar{\alpha}_n) \leqq 4\epsilon^2\|f\|^2 \qquad \text{for} \qquad n \geqq 1/\epsilon^2 w_\epsilon.$$

It was our intention to make the estimation as exact as possible.

Our construction will furnish *all* the primitive invariant sets occurring in f. This is readily seen in the following manner. It will furnish certain, let us say l' of the l mutually orthogonal primitive harmonic sets of order U. The remaining $l - l'$ cannot occur in f or they contribute nothing to the total magnitude of f, as is obvious from Bessel's inequality according to which the sum of all contributions cannot exceed $\|f\|^2$ whereas the contributions *which are furnished by our construction* alone give the whole amount $\|f\|^2$. Hence we may formulate our final result thus:

THEOREM 4. *We gave a construction yielding all the primitive invariant sets occurring in f, in a definite normalization and order depending on f.* (The order is such that the weights with which the successive sets occur in f form a monotonely decreasing sequence.) *The relation of completeness is valid; it can be supplemented by the explicit estimation of the remainder* (28) *if one keeps to the normalization and order originating from the construction.*

The *expansion theorem* now is an immediate consequence. We consider a function $g(P)$ of the form

$$g(P) = \int_s f(s^{-1}P)\, x(s).$$

Writing again

$$f(P) = \{e(P) + e_1(P) + \cdots + e_{p-1}(P)\} + f_p(P)$$

one has, according to Schwarz's inequality,

$$\left| \int_s f_p(s^{-1}P)\, x(s) \right|^2 \leqq \int_P |f_p(P)|^2 \cdot \int_s |x(s)|^2$$

and the remainder as $\int_s f_p(s^{-1}P)\, x(s)$ tends therefore to zero with increasing p in a manner that can be estimated explicitly. For the individual contribution we get from (22):

$$\int_s e(s^{-1}P)\, x(s) = \sum_{\mu=1}^{h} \beta_\mu \varphi_\mu(P) \qquad \text{with} \qquad \beta_\mu = \int_s \psi_\mu(s)\, x(s).$$

β_μ coincides with the Fourier coefficient of $g(P)$ with respect to $\varphi_\mu(P)$:

$$\int_P g(P)\, \bar{\varphi}_\mu(P) = \int_s \int_P f(s^{-1}P)\, \bar{\varphi}_\mu(P)\, x(s) = \int_s \psi_\mu(s)\, x(s).$$

THEOREM 5. *The Fourier series*

$$\beta_1\varphi_1(P) + \beta_2\varphi_2(P) + \cdots, \qquad \beta_i = \int g(P)\, \bar{\varphi}_i(P)$$

of a function $g(P)$ of the form

$$g(P) = \int_s f(s^{-1} P)\, x(s)$$

converges uniformly to $g(P)$.

Only the primitive sets of functions actually occurring *in* f appear in this expansion. They are ordered by decreasing weight in f, and the summation should be performed "by sections," without separating the terms of the individual primitive sets of functions.

One obtains the *theorem of approximation* in choosing $x(s)$ as a function which is $\neq 0$ only in the immediate neighborhood of the unit element:

THEOREM 6. *$f(P)$ can be approximated uniformly by a finite sum of harmonics occurring in f.*

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

A SOLUTION OF THE PRINCIPAL PROBLEM IN THE THEORY OF RIEMANN MATRICES[1]

BY A. ADRIAN ALBERT

(Received March 28, 1934)

1. **Introduction.** The principal problem in the theory of Riemann matrices is that of determining the multiplication algebra D of any pure Riemann matrix ω. This problem is of great importance in the transcendental theory of algebraic geometry,[2] and its solution provides a solution of the many related algebraic geometric questions from which it arose.

Algebra D is a normal division algebra of degree n (order n^2) over its centrum K. The centrum K is an algebraic field $R(\rho)$ of all rational functions with rational coefficients of an abstract root ρ of an irreducible rational equation $f(r) = 0$. We call ω a Riemann matrix of the first or second kind according as $f(r) = 0$ does or does not have all real roots.

If ω is of the first kind, then[3] $n = 1, 2$; I have determined[4] the resulting algebras D, and have proved[4] that there exist pure Riemann matrices ω with any one of the so determined algebras D as multiplication algebra. There remains the case of matrices ω of the second kind.

The remaining more difficult problem will be solved in the present paper. I shall first define a type of algebra $\mathfrak{A}$, shall prove the existence of Riemann matrices ω with $\mathfrak{A}$ as multiplication algebra, and shall finally show that the multiplication algebra D of any pure Riemann matrix ω of the second kind is an algebra $\mathfrak{A}$.

2. **An elementary existence theorem.** Let F_0 be a real algebraic number field, $\mu_1 < 0$ in F_0, $\mu_2 < 0$ in F_0, and let K_0 be an algebraic number field containing F_0, $\mu_1^{\frac{1}{2}}$, $\mu_2^{\frac{1}{2}}$. Consider matrices[5]

$$(1) \qquad \begin{aligned} \tau_1 &= \| I(g_1) \quad \pi_1 \|, & \tau_2 &= \| -\pi_1' \quad -I(g_2) \| \\ \tau_3 &= \| I(g_1) \quad \pi_2 \|, & \tau_4 &= \| -\pi_2' \quad -I(g_2) \| \end{aligned}$$

---

[1] Presented to the Society March 30, 1934.

[2] For references to this theory see S. Lefschetz's report in the Report of the Committee on Rational Transformations, National Research Council Bulletin No. 63, Chapters XV, XVI, XVII, pp. 310–393.

[3] These Annals, vol. 33 (1932), pp. 311–318.

[4] In my paper *On the construction of Riemann matrices* II, which will appear in these Annals.

[5] We use the notation $I(g)$ for an identity matrix of g-rows. We also use 0 to represent zero or zero matrices.

where π_1 and π_2 are g_1-rowed and g_2-columned matrices of complex elements. Then if

$$g = g_1 + g_2, \qquad g_1 > 1,\ g_2 > 1, \tag{2}$$

and

$$\Gamma = \begin{Vmatrix} I(g_1) & 0 \\ 0 & -I(g_2) \end{Vmatrix}, \tag{3}$$

we have $\Gamma' = \Gamma$,

$$\tau_1 \Gamma \tau_2' = \| I(g_1) \quad \pi_1 \| \cdot \begin{Vmatrix} I(g_1) & 0 \\ 0 & -I(g_2) \end{Vmatrix} \cdot \begin{Vmatrix} -\pi_1 \\ -I(g_2) \end{Vmatrix} = -\pi_1 + \pi_1 = 0,$$

$$\tau_1 \Gamma \bar{\tau}_1' = \| I(g_1) \quad \pi_1 \| \cdot \begin{Vmatrix} I(g_1) & 0 \\ 0 & -I(g_2) \end{Vmatrix} \cdot \begin{Vmatrix} I(g_1) \\ \bar{\pi}_1' \end{Vmatrix} = I(g_1) - \pi_1 \bar{\pi}_1', \tag{4}$$

$$-\tau_2 \Gamma \bar{\tau}_2' = \| \pi_1' \quad I(g_2) \| \cdot \begin{Vmatrix} I(g_1) & 0 \\ 0 & -I(g_2) \end{Vmatrix} \cdot \begin{Vmatrix} -\bar{\pi}_1 \\ -I(g_2) \end{Vmatrix} = I(g_2) - \pi_1' \bar{\pi}_1.$$

Similarly,

$$\tau_3 \Gamma \tau_4' = 0, \qquad \tau_3 \Gamma \bar{\tau}_3' = I(g_1) - \pi_2 \bar{\pi}_2', \qquad -\tau_4 \Gamma \bar{\tau}_4' = I(g_2) - \pi_2' \bar{\pi}_2. \tag{5}$$

We write

$$\pi_1 = \| \pi_{1jk} \|, \qquad \pi_2 = \| \pi_{2jk} \| \qquad (j = 1, \cdots, g_1;\ k = 1, \cdots, g_2). \tag{6}$$

It is known[6] that there exist complex algebraic numbers π_{1jk}, π_{2jk} which are sufficiently small in absolute value such that

$$1,\ \pi_{1jk},\ \pi_{2jk},\ \pi_{1jk}\pi_{1st},\ \pi_{1jk}\pi_{2st} \qquad (j, s = 1, \cdots, g_1;\ k, t = 1, \cdots, g_2) \tag{7}$$

are linearly independent with respect to K_0 and the matrices $\tau_1 \Gamma \bar{\tau}_1'$, $-\tau_2 \Gamma \bar{\tau}_2'$, $\tau_3 \Gamma \bar{\tau}_3'$, $-\tau_4 \Gamma \bar{\tau}_4'$ are all positive definite.

Let B be any g-rowed square matrix with elements in K_0, and suppose that there exists a complex matrix ρ such that $\rho\tau_2 = \tau_1 B$. Then

$$\tau_1 B \Gamma \tau_1' = \rho \tau_2 \Gamma \tau_1' = \rho(\tau_1 \Gamma \tau_2')' = 0.$$

Define $B\Gamma = E$,

$$E = \begin{Vmatrix} E_1 & E_2 \\ E_3 & E_4 \end{Vmatrix}, \quad \tau_1 E \tau_1' = E_1 + \pi_1 E_3 + E_2 \pi_1' + \pi_1 E_4 \pi_1' = 0, \tag{8}$$

where

---

[6] See my paper *On the construction of Riemann matrices* I, these Annals, vol. 35 (1934), pp. 1-28.

$$(9)\quad E_1 = \| E_{1jk} \|, \qquad E_2 = \| E_{2jt} \|, \qquad E_3 = \| E_{3ks} \|, \qquad E_4 = \| E_{4kt} \|,$$
$$(j, s = 1, \cdots, g_1;\ k, t = 1, \cdots, g_2).$$

The element in the j^{th} row and s^{th} column of $\tau_1 E \tau_1'$ is

$$(10)\qquad E_{1js} + \sum_k \pi_{1jk} E_{3ks} + \sum_t E_{2jt} \pi_{1st} + \sum_{k,t} \pi_{1jk} E_{4kt} \pi_{1st} = 0.$$

The linear independence of (7) in K_0 and (10) imply that

$$E_1 = E_4 = 0, \qquad \sum_k \pi_{1jk} E_{3ks} + \sum_t E_{2jt} \pi_{1st} = 0.$$

Since $g_1 > 1$ we may take $j \neq s$ and obtain $E_{3ks} = E_{2jt} = 0$ and hence $E_2 = E_3 = 0$. But then $E = 0$ so that evidently $B = 0$.

Similarly, if B has elements in K_0, and is such that either $\rho\tau_1 = \tau_2 B$ or $\rho\tau_3 = \tau_4 B$ or $\rho\tau_4 = \tau_3 B$ then $B = 0$.

We next assume that $\rho\tau_1 = \tau_1 E$ for complex ρ and where E is given by (8), (9). Then

$$(11)\qquad \rho = E_1 + \pi_1 E_3, \qquad \rho\pi_1 = E_1 \pi_1 + \pi_1 E_3 \pi_1 = E_2 + \pi E_4,$$

from which we obtain

$$(12)\qquad \sum_s E_{1js} \pi_{1st} + \sum_{k,s} \pi_{1jk} E_{3ks} \pi_{1st} = E_{2jt} + \sum_k \pi_{1jk} E_{4kt}.$$

But (7), (12) imply that $E_3 = E_2 = 0$, $E_{1js} = E_{4kt} = 0$ unless $j = s$, $t = k$, and $E_{1jj} = E_{4tt}$, so that E is a scalar matrix.

Similarly $\rho\tau_2 = \tau_2 E$, $\rho\tau_3 = \tau_3 E$, $\rho\tau_4 = \tau_4 E$; each imply that E is scalar.

Let $F_1, \cdots, F_t$ be real algebraic fields contained in F_0, let μ_j in F_j be negative, and let K_0 contain F_0 and $\mu_j^{\frac{1}{2}}$. Let

$$(13)\qquad \Gamma_j = \begin{Vmatrix} \lambda_{j1} I(g_0) & & 0 \\ & \ddots & \\ 0 & & \lambda_{jm} I(g_0) \end{Vmatrix} \cdot \Delta, \qquad \Delta = \begin{Vmatrix} \beta & & 0 \\ & \ddots & \\ 0 & & \beta \end{Vmatrix},$$

be a matrix of mg_0 rows where $m > 1$, $g_0 > 1$ and λ_{jk} are in F_j. Moreover, let β be a g_0-rowed diagonal matrix whose diagonal elements consist of g_{01} elements $+1$ and g_{02} elements -1,

$$(14)\qquad g_0 = g_{01} + g_{02}, \qquad g_{01} g_{02} \neq 0.$$

Assume that $\lambda_{jk} > 0$ and let K_0 contain $(\lambda_{jk})^{\frac{1}{2}}$. Then evidently there exists a non-singular matrix N_j with (real) elements in K_0 such that

$$(15)\qquad N_j \Gamma_j N_j' = \Gamma,$$

when we write

$$g = mg_0, \qquad g_1 = mg_{01}, \qquad g_2 = mg_{02},$$

and take Γ to be the matrix given by (3). We now let

$$\tau^{(11)} = \tau_1 N_1, \qquad \tau^{(12)} = \tau_2 N_1, \qquad \tau^{(j1)} = \tau_3 N_j, \qquad \tau^{(j2)} = \tau_4 N_j, \tag{16}$$
$$(j = 2, \cdots, t).$$

Evidently $\tau^{(j1)} \Gamma_j \tau^{(j2)'} = \tau_1 \Gamma \tau_2'$ or $\tau_3 \Gamma \tau_4'$ and is zero. Moreover $\tau^{(j1)} \Gamma_j \overline{\tau^{(j1)'}}$, $-\tau^{(j2)} \Gamma_j \overline{\tau^{(j2)'}}$ are positive-definite Hermitian by use of (15), the reality of N_j, and the fact that $\tau_1 \Gamma \bar{\tau}_1', -\tau_2 \Gamma \bar{\tau}_2', \tau_3 \Gamma \bar{\tau}_3', -\tau_4 \Gamma \bar{\tau}_4'$ have been chosen positive definite Hermitian.

If $\rho\tau^{(11)} = \tau^{(12)} E$ then $\rho\tau_1 = \tau_2 N_j E N_j^{-1}$ where $N_j E N_j^{-j}$ has elements in K_0. But then $N_j E N_j^{-1} = 0$ so that $E = 0$. Similarly $\rho\tau^{(11)} = \tau^{(j1)} E$ implies that $E = 0$ when $j \neq 1$, while $\rho\tau^{(11)} = \tau^{(11)} E$ is equivalent to $\rho\tau_1 = \tau N_1 E N_1^{-1}$ so that $N_1 E N_1^{-1} = eI(g)$ with e in K_0. Hence $E = eI(g)$ is scalar.

Similarly $\rho\tau^{(j1)} = \tau^{(j1)} E$ or $\rho\tau^{(j2)} = \tau^{(j2)} E$ implies that E is a scalar matrix. We have proved an elementary result which is of fundamental importance for our later construction of Riemann matrices.

THEOREM 1. *Let $F_1, \cdots, F_t$ be real algebraic number fields, let $\mu_j < 0$ in F_j, and let K_0 be the composite of $F_j(\mu_j^{\frac{1}{2}})$. Write*

$$\Gamma_j = \begin{Vmatrix} \lambda_{j1}\beta & & 0 \\ & \ddots & \\ 0 & & \lambda_{jm}\beta \end{Vmatrix}, \ \lambda_{jk} > 0 \text{ in } F_j, \qquad (j = 1, \cdots, t) \tag{17}$$

where β is a diagonal matrix with g_{01} diagonal elements 1, g_{02} diagonal elements -1. Assume that $m > 1$, $g_{01} g_{02} \neq 0$, $g_0 = g_{01} + g_{02}$, $g = mg_0$, $g_1 = mg_{01}$, $g_2 = mg_{02}$. Then there exist g-columned matrices $\tau^{(j1)}$, $\tau^{(j2)}$ of g_1 and g_2 rows respectively such that

$$\tau^{(j1)} \Gamma_j \tau^{(j2)'} = 0,$$

and

$$\tau^{(j1)} \Gamma_j \overline{\tau^{(j1)'}}, \qquad -\tau^{(j2)} \Gamma_j \overline{\tau^{(j2)'}}$$

are positive definite Hermitian matrices. Finally $\tau^{(j1)}$ and $\tau^{(j2)}$ may be so chosen that, if B is any matrix with elements in K_0 and ρ is any complex matrix with the property

$$\rho\tau^{(11)} = \tau^{(12)} B, \text{ or } \rho\tau^{(12)} = \tau^{(11)} B, \text{ or } \rho\tau^{(11)} = \tau^{(j1)} B, \text{ or } \rho\tau^{(11)} = \tau^{(j2)} B,$$
$$(j = 2, \cdots, t),$$

then $B = 0$, while each of $\rho\tau^{(j1)} = \tau^{(j1)} B$ and $\rho\tau^{(j2)} = \tau^{(j2)} B$ $(j = 1, \cdots, t)$ implies that B is a scalar matrix.

3. **A certain type of algebra.** Let R be the field of all rational numbers,

$$f(\lambda) \equiv \lambda^t + a_1 \lambda^{t-1} + \cdots + a_t = 0, \qquad (a_i \text{ rational}), \tag{18}$$

be irreducible in R and have all real roots $\sigma_1, \cdots, \sigma_t$. Then the field $F = R(\mathfrak{s})$ of all rational functions with rational coefficients of an abstract root $\mathfrak{s}$ of $f(\lambda) = 0$ is called a total real algebraic extension of R.

A quantity $\nu = \nu(\mathfrak{s})$ is called total positive if $\nu(\sigma_j) > 0, j = 1, \cdots, t$. Similarly $\mu = \mu(\mathfrak{s})$ is total negative if $\mu(\sigma_j) < 0$ and, in this latter case, the field

$$K = F(q), \qquad q^2 = \mu = \mu(\mathfrak{s}) \text{ in } F, \tag{19}$$

is called a total pure imaginary extension of F. We then write

$$F_{0j} = R(\sigma_j), \qquad K_j = F_{0j}(\mu_j^{\frac{1}{2}}), \qquad \mu_j = \mu(\sigma_j). \tag{20}$$

Consider an equation

$$\phi(\lambda) \equiv \lambda^m + \alpha_1\lambda^{m-1} + \cdots + \alpha_m = 0 \qquad (\alpha_i \text{ in } F), \tag{21}$$

which has the cyclic group for F and hence roots

$$\theta^m(x) = x, \theta(x), \theta^2(x), \cdots, \theta^{m-1}(x) \tag{22}$$

in a cyclic field $Z = F(x)$. Moreover let $Z = F(x)$ be total real. Then

$$\phi_j(\lambda) \equiv \lambda^m + \alpha_1(\sigma_j)\lambda^{m-1} + \cdots + \alpha_m(\sigma_j) = 0 \tag{23}$$

has the cyclic group for F_{0j} and roots

$$\xi_j, \qquad \theta_j(\xi_j), \cdots, \theta_j^{m-1}(\xi_j) \tag{24}$$

all real. Write

$$\theta_j^{e-1}(\xi_j) = \xi_{je} \qquad (j = 1, \cdots, t; e = 1, \cdots, m). \tag{25}$$

We shall use the well known trivial

LEMMA 1. *Let F be any non-modular field, Z be algebraic of degree m over F, K be quadratic over F. Then the composite $W = (Z, K)$ over K has degree $\frac{1}{2}(m)$ or m according as Z does or does not have a sub-field equivalent to K.*

The total real cyclic field Z evidently does not have a quadratic subfield equivalent to the total pure imaginary field K so that $W = (Z, K)$ is cyclic of degree m over K. We then define a cyclic algebra $\mathfrak{A}$ with basis

$$x^i y^j \qquad (i, j = 0, 1, \cdots, m-1) \tag{26}$$

and multiplication table

$$y^i x = \theta^i(x) y^i, \qquad y^m = \gamma \qquad (i = 0, 1, \cdots), \tag{27}$$

where γ in K has the form

$$\gamma = \gamma_1 + \gamma_2 q \qquad (\gamma_1, \gamma_2 \text{ in } F). \tag{28}$$

Moreover, let there exist a total positive quantity d in $F(x)$ such that

$$\gamma_1^2 - \gamma_2^2\mu = d(x) \cdot d[\theta(x)] \cdots d[\theta^{m-1}(x)] \equiv N_{Z/F}(d). \tag{29}$$

We shall now prove the fundamental existence theorem.

THEOREM 2. *There exist Riemann matrices*[7] ω *of genus*

$$g_3 = tm^2 g_0, \qquad g_0 > 1 \tag{33}$$

with $\mathfrak{A}$ *as multiplication algebra.*

We take

$$\omega = \left\| \begin{matrix} \omega^{(1)} & & & 0 \\ & \omega^{(2)} & & \\ & & \ddots & \\ 0 & & & \omega^{(t)} \end{matrix} \right\| \cdot V_s, \qquad \omega^{(j)} = \left\| \begin{matrix} \omega^{(j1)} & 0 \\ 0 & \omega^{(j2)} \end{matrix} \right\| \cdot W_j,$$

$$\omega^{(j1)} = \left\| \begin{matrix} \tau_1^{(j1)} & & & 0 \\ & \tau_2^{(j1)} & & \\ & & \ddots & \\ 0 & & & \tau_m^{(j1)} \end{matrix} \right\| \cdot V_x^{(j)}, \qquad \omega^{(j2)} = \left\| \begin{matrix} \tau_1^{(j2)} & & & 0 \\ & \tau_2^{(j2)} & & \\ & & \ddots & \\ 0 & & & \tau_m^{(j2)} \end{matrix} \right\| \cdot V_x^{(j)}, \tag{34}$$

where

$$V_s = \left\| I(2m^2 g_0)\, \sigma_j^{k-1} \right\|, \qquad W_j = \left\| \begin{matrix} I(m^2 g_0) & \mu_j^{\frac{1}{2}} I(m^2 g_0) \\ I(m^2 g_0) & -\mu_j^{\frac{1}{2}} I(m^2 g_0) \end{matrix} \right\|, \tag{35}$$

$$V_x^{(j)} = \left\| I(mg_0)\, \xi_{j\,e}^{f-1} \right\|, \qquad (j, k = 1, \cdots, t;\ e, f = 1, \cdots, m).$$

We also write

$$\Lambda_j = \left\| \begin{matrix} 0 & I(g_0) & 0 & \cdots & 0 \\ 0 & 0 & I(g_0) & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & I(g_0) \\ \gamma_j I(g_0) & 0 & 0 & \cdots & 0 \end{matrix} \right\|, \tag{36}$$

where $\gamma_j = \gamma_1(\sigma_j) + \mu_j^{\frac{1}{2}} \gamma_2(\sigma_j)$. Finally write

$$C^{(j)}(\xi_1) = \left\| \begin{matrix} \beta_1^{(j)} & & & 0 \\ & \beta_2^{(j)} & & \\ & & \ddots & \\ 0 & & & \beta_m^{(j)} \end{matrix} \right\| \cdot (-\mu_j^{\frac{1}{2}}). \tag{37}$$

Then if $d_j = d(\xi_j)$ we have

$$\Lambda_j C^{(j)}(\xi_{j2})\, \overline{\Lambda}_j' = d_j C^{(j)}(\xi_{j1}), \qquad (j = 1, \cdots, t), \tag{38}$$

if, and only if,

$$\beta_{e+1}^{(j)}(\xi_{j2}) = d_j \beta_e^{(j)} \quad (e = 1, \cdots, m-1), \qquad \beta_1^{(j)}(\xi_{j2})\, \gamma_j \bar{\gamma}_j = d_j \beta_m^{(j)}. \tag{39}$$

The first set of these equations evidently determines $\beta_2, \cdots, \beta_m$ in terms of β_1.

---

[7] These matrices are pure if, and only if, $\mathfrak{A}$ is a division algebra.

They imply

$$(40)\qquad \beta_2^{(j)}(\xi_{j2}) = d_j\beta_1, \qquad \beta_3^{(j)}(\xi_{j3}) = d_j(\xi_{j2})\, d_j(\xi_{j1})\, \beta_1^{(j)}, \cdots, \beta_m^{(j)}(\xi_{j,m-1}) = d_j(\xi_{j,m-2}) \cdots d_j(\xi_{j1})\, \beta_1^{(j)},$$

so that

$$(41)\qquad d_j\, \beta_m^{(j)}\ (\xi_{jm}) = d_j\, \beta_m^{(j)} = N_{Z_j/F_{0j}}\,(d_j) \cdot \beta_1^{(j)}\ (\xi_{j2}).$$

Equation (29) implies that $N_{Z_j/F_{0j}}\,(d_j) = \gamma_j\,\bar{\gamma}_j$ and (39) are satisfied. We now take $\beta_1^{(j)} = \beta$ to be a g_0-rowed diagonal matrix with g_{01} diagonal elements unity, g_{02} diagonal elements -1, and $g_{01}g_{02} \neq 0$, and we may write

$$(42)\qquad C^{(j)}(\xi_{j1}) = -\,\mu_j^{\frac{1}{2}}\,\Gamma_j,$$

where Γ_j has the form (17) since $\lambda_{jk} = d_j(\xi_{jk}) > 0$.

We next take

$$(43)\qquad \tau_e^{(j1)} = \tau^{(j1)}\,\Lambda_j^{e-1}, \qquad \tau_e^{(j2)} = \tau^{(j2)}\,\overline{\Lambda}_j^{e-1}$$

where $\tau^{(j1)}$ and $\tau^{(j2)}$ are chosen as in Theorem 1 with $F_j = F_{j0}(\xi_1)$. Then

$$(44)\qquad \tau^{(j1)}\, C^{(j)}\,(\xi_{j1})\, \tau^{(j2)\prime} = 0, \qquad (-1)^{\frac{1}{2}}\,\tau^{(j1)}\, C^{(j)}(\xi_{j1})\, \overline{\tau^{(j1)\prime}}, \qquad (-1)^{\frac{1}{2}}\,\tau^{(j2)}\, \overline{C^{(j)}(\xi_{j1})}\ \overline{\tau^{(j2)\prime}},$$

are positive definite Hermitian since $(-1)^{\frac{1}{2}}\, C^{(j)}(\xi_{j1}) = -\,(-1)^{\frac{1}{2}}\,\mu_j^{\frac{1}{2}}\,\Gamma_j = (-\mu_j)^{\frac{1}{2}}\,\Gamma_j$ where $-\mu_j > 0$.

By (38) we have

$$(45)\qquad \Lambda_j\, C^{(j)}(\xi_{j,e+1})\, \overline{\Lambda}_j' = d_j(\xi_{je})\, C^{(j)}(\xi_{je}),$$

so that

$$(46)\qquad \Lambda_j^{e-1}\, C^{(j)}(\xi_{je})\, (\overline{\Lambda}_j')^{e-1} = \nu_{je}\, C^{(j)}(\xi_j), \qquad (e = 2, \cdots, m),$$

where $\nu_{je} > 0$. Hence

$$\tau_e^{(j1)}\, C^{(j)}(\xi_{je})\, \tau_e^{(j2)\,\prime} = 0, \qquad (-1)^{\frac{1}{2}}\,\tau_e^{(j1)}\, C^{(j)}(\xi_{je})\, \overline{\tau_e^{(j1)\,\prime}}, \qquad (-1)^{\frac{1}{2}}\,\tau_e^{(j2)}\, \overline{C^{(j)}(\xi_{je})}\, \overline{\tau_e^{(j2)}},$$

are positive definite Hermitian matrices. But then, as I have shown,[8] ω is a Riemann matrix.

The algebra $\mathfrak{B}$ of all the multiplications of ω which are commutative with s and g, has $\mathfrak{A}$ as normal simple sub-algebra, and hence[9] differs from $\mathfrak{A}$ if, and only if, $\mathfrak{B}$ contains a multiplication b not in $R(s, q)$ but such that $bx = xb$, $by = yb$. Hence it is sufficient to prove that *every multiplication b of ω has the property*

---

[8] See my paper *On the structure of matrices with any normal division algebra of multiplications*, these Annals, vol. 32 (1931), pp. 131–148. The above result is implied by the proof of Theorem 11 with P_e of Theorem 11 our above matrix Λ_j^{e-1}.

[9] For the well known theorem of J. H. M. Wedderburn states that $\mathfrak{B}$ is the direct product of $\mathfrak{A}$ and another algebra $\mathfrak{C}$ over $R(s, q)$ and $\mathfrak{A} \neq \mathfrak{B}$ if, and only if, $\mathfrak{C} \neq R(s, q)$.

$bs = sb$, $bg = gb$, *and when also* $bx = xb$, $by = yb$, *then* b *is in* $R(s, g)$, in order to prove that $\mathfrak{A}$ is the multiplication algebra of ω.

I have proved[10] that every multiplication of ω is commutative with s if the equation $\rho_j\omega^{(1)} = \omega^{(j)}B(\sigma_j, \sigma_1)$ implies that $B(\sigma_j, \sigma_1)$ is a zero matrix. If we write

$$W_jB(\sigma_j, \sigma_1)\, W_1^{-1} = \left\| \begin{matrix} B(\mu_j^{\frac{1}{2}}, \mu_1^{\frac{1}{2}}) & B(\mu_j^{\frac{1}{2}}, -\mu_1^{\frac{1}{2}}) \\ B(-\mu_j^{\frac{1}{2}}, \mu_1^{\frac{1}{2}}) & B(-\mu_j^{\frac{1}{2}}, -\mu_1^{\frac{1}{2}}) \end{matrix} \right\| \tag{47}$$

then $\rho_j\omega^{(1)} = \omega^{(j)}B(\sigma_j, \sigma_1)$ implies that

$$\rho_{j1}\,\omega^{(11)} = \omega^{(j1)}B(\mu_j^{\frac{1}{2}}, \rho_1^{\frac{1}{2}}), \qquad \rho_{j2}\omega^{(12)} = \omega^{(j1)}B(\mu_j^{\frac{1}{2}}, -\mu_j^{\frac{1}{2}}), \tag{48}$$

for complex ρ_{j1}, ρ_{j2}. But then also

$$\rho^{(j1)}\,\tau^{(11)} = \tau^{(j1)}B_{j1}, \qquad \rho^{(j2)}\,\tau^{(11)} = \tau^{(j1)}B_{j2} \tag{49}$$

where $\rho^{(j1)}$, $\rho^{(j2)}$ are complex matrices, $B_{j1} = B_{j1}(\sigma_j, \sigma_1, \xi_j, \xi_1, \mu_j^{\frac{1}{2}}, \mu_1^{\frac{1}{2}})$, $B_{j2} = B_{j1}(\sigma_j, \sigma_1, \xi_j, \xi_1, \mu_j^{\frac{1}{2}}, -\mu_1^{\frac{1}{2}})$ and $B(\sigma_j, \sigma_1) = 0$ if, and only if, $B_{j1} = B_{j2} = 0$. By Theorem 1 equations (49) imply that $B_{j1} = B_{j2} = 0$, and hence that every multiplication of ω is commutative with s. Thus every multiplication $\beta\omega = \omega B$ of ω defines a multiplication b^j of the Riemann matrix $\omega^{(j)}$ over F_{0j} given by

$$\beta_j\,\omega^{(j)} = \omega^{(j)}\,B_j \tag{50}$$

where $B_j = B_j\,(\sigma_j)$. Then b is necessarily commutative with q if we can prove b^j commutative with q_j. But we have shown in Theorem 1 that if E has elements in $R(\sigma_1, \cdots, \sigma_t; \xi_1, \cdots, \xi_t, \mu_1^{\frac{1}{2}}, \cdots, \mu_t^{\frac{1}{2}}) = K_0$ and is such that $\rho\,\tau^{(11)} = \tau^{(12)}E$ or $\rho\,\tau^{(21)} = \tau^{(11)}E$ then $E = 0$. I have also proved[11] that then every multiplication b_j of $\omega^{(j)}$ is commutative with q_j. Hence every multiplication b of ω is commutative with q.

Finally let b be any multiplication of ω, so that necessarily $bs = sb$, $bq = qb$, and, if b is not in $\mathfrak{A}$, then we may also assume $bx = sb$, $by = yb$. If b is given by $\beta\omega = \omega B$, we have

$$V_s\,BV_s^{-1} = \left\| \begin{matrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_t \end{matrix} \right\|, \quad W_j\,B_j\,W_s^{-1} = \left\| \begin{matrix} B_{j1} & 0 \\ 0 & B_{j2} \end{matrix} \right\|, \tag{51}$$

and

$$V_x^{(j)}\,B_{j1}\,V_x^{(j)-1} = \left\| \begin{matrix} B_{j11} & & 0 \\ & \ddots & \\ 0 & & B_{jm1} \end{matrix} \right\|, \tag{52}$$

[10] *On the construction of Riemann matrices* I, loc. cit.

[11] On the structure of matrices, etc., loc. cit. Theorem 11.

since b is commutative with s, q, x. The matrix $B_{jk1} = E(\sigma_j, \mu_j^{\frac{1}{2}}, \xi_{jk})$ is non-scalar if b is not in $R(s, q)$. But $\beta\omega = \omega B$ implies that $\beta_{jk}\, \tau^{(11)} = \tau^{(11)}\, B_{j11}$. By Theorem 1, B_{j11} is scalar. Hence so are all the B_{jm1}, a contradiction.

We have proved that the Riemann matrix ω of (34) has $\mathfrak{A}$ as multiplication algebra and hence Theorem 2.

4. **The arithmetic theory.**[12] Let

$$\phi(\lambda) \equiv \lambda^a + b_1 \lambda^{a-1} + \cdots + b_a = 0 \qquad (b_j \text{ in } R) \tag{53}$$

be irreducible in the field R of all rational numbers and have complex roots $r_1, \cdots, r_a$. If r is an arbitrary complex root of $\phi(\lambda) = 0$ we consider a field $F = R(r)$, an algebraic field of finite degree over R.

The roots $r_1, \cdots, r_a$ define infinite prime spots $p_1, \cdots, p_a$ of F and we define F_{p_j} to be the field of all real or all complex numbers according as r_j is real or imaginary.

Let Z be a galois extension of F. We may define the infinite prime spots $\mathfrak{P}$ of Z as for F and it is evident that every $Z_{\mathfrak{P}}$ contains an F_p. If F_p is equivalent to the field of all complex numbers, then so evidently is $Z_{\mathfrak{P}}$, and we say that Z has p-degree one. If F_p is equivalent to the field of all real numbers, then all $Z_{\mathfrak{P}}$ containing F_p are equivalent (since Z is a galois extension of F), and Z is said to have p-degree one or two according as $Z_{\mathfrak{P}}$ is or is not real.

The finite prime spots of F are its prime ideals p and we define the p-adic extensions F_p of F. If Z is galois over F, then every prime ideal $\mathfrak{P}$ of Z divides a prime ideal p of F. Moreover we have[13]

LEMMA 2. *Let Z be a galois extension of F, p a prime ideal of F, $\mathfrak{P}$ a prime divisor of p in Z. Then*

$$Z_{\mathfrak{P}} = (Z, F_p).$$

Since $Z_{\mathfrak{P}}$ is thus the composite of Z and F_p, its degree over F_p is independent of $\mathfrak{P}$ and is called the p-degree of F.

If A is any normal simple algebra over F and we extend F to F_p, the corresponding algebra A_p over F_p is also normal simple. The index of A is the degree of its normal division algebra factor, and is divisible by the index of A_p, the p-index of A. If Z is any extension of F, then Z is said to split A if A_Z over Z is a total matric algebra. When A and Z both have degree m, the field Z splits A if, and only if, A has a sub-field equivalent to Z. Every A over F is now known to be a cyclic algebra so we may use the following results of Hasse.

---

[12] Cf. Hasse's paper in the Transactions of the American Mathematical Society, vol. 34 (1932), pp. 171–214, for the division algebra theory and references to the arithmetic theory of algebraic number fields.

[13] Cf. H. Hasse's Klassenkörpertheorie for a proof of this theorem.

LEMMA 3. *A cyclic field Z over F is a splitting field for A if, and only if, the p-degrees of Z are multiples of the p-indices of A for every prime spot p of F.*

LEMMA 4. *The p-indices of A are unity except for a finite number of prime spots p of F.*

We shall also assume a corollary of Grunwald's important existence theorem.[14]

LEMMA 5. *Let $p_1, \cdots, p_r$ be prime ideals of F, n a positive integer. Then there exist cyclic fields Z of degree n over F with p_i-degrees n $(i = 1, \cdots, r)$.*

Finally we assume the trivial well-known algebraic theorem Lemma 1.

5. **Algebras over K.** Let F be a total real algebraic number field, let K be a total pure imaginary extension of F. If π is any infinite prime spot of K, p a corresponding infinite prime spot of F, it is evident that K_π is the field of all complex numbers, F_p the field of all real numbers. If A is a normal simple algebra over K, then A has π-index unity for every infinite prime spot π of K. Hence we need only consider the finite spots or prime ideals of K.

Let D be a normal division algebra of degree n over K, and $\pi_1, \cdots, \pi_r$ be the prime spots (all finite prime ideals) for which the π-index of D is not unity. Let $p_1, \cdots, p_r$ be the corresponding prime ideals in F divisible by the π_i.

If n is odd, let $A = D$, $m = n$. If

$$n = 2n_1 = 2^\nu n_0 \qquad (n_0 \text{ odd}, \nu > 0), \tag{54}$$

we write $m = 2n$, $A = D \times M_2$ where M_2 is a total matric algebra of degree 2. In either case there exists a cyclic field Z of degree m over F such that Z has p_i-degrees m. We shall prove

LEMMA 6. *The composite $W = (Z, K)$ is a cyclic splitting field of degree m of A, and hence is equivalent to a sub-field of A.*

The field W has degree m over K if, and only if, Z has no quadratic sub-field equivalent to K. If n is odd this is evidently true and W is obviously cyclic of degree $m = n$ over K. If n is even, then $m = 4n_1$, Z has a unique cyclic sub-field Z_4 of degree four over F, and a unique quadratic sub-field Z_2 in Z_4. But it is well known[15] that $Z_2 = F(u)$, $u^2 = 1 + \epsilon^2$ with ϵ in F. Since F is total real, $1 + \epsilon^2$ is total positive and Z_2 is total real. But then K is not equivalent to Z_2 and W has degree m over K and obviously cyclic.

The field $W\mathfrak{P}_i$ is the composite of W and K_{π_i} by Lemma 2. But W is the composite of Z and K contained in K_{π_i} so that $W\mathfrak{P}_i$ is the composite of Z and $K_{\pi_i} = (K, F_{p_i})$. It is therefore evident that $W\mathfrak{P}_i$ is the composite of Z_{P_i} over F_{p_i} and K_{π_i} over F_{p_i}.

The field Z_{P_i} is assumed to have degree m while K_{π_i} has degree one or two.

---

[14] Journal für Mathematik, vol. 169 (1933), pp. 103–107. We have here used his Theorem 3 with $e_k = 1$, $f_k = m$, so that the p_k remain prime ideals in Z and hence have relative order 1 and relative degree m. But then Z has p_k degree m since in fact the p_k-degree of Z is $e_k \cdot f_k$.

[15] Cf. my paper on *Cyclic fields of degree eight*, transactions of the American Mathematical Society, vol. 35 (1933), pp. 949–964.

By Lemma 1 $W\mathfrak{P}_i$ has degree m or $\frac{m}{2}$ over K according as Z_{P_i} does not or does have a quadratic sub-field equivalent to K_{π_i}. When n is odd, $m = n$ and Z_{P_i} has no quadratic sub-field. In this case W has π_i-degree m and splits $A = D$ by Lemma 3. When n is even, then $m = 2n$ so that W has π_i-degree at least n and W is again a splitting field of D.

LEMMA 7. *The field W contains a total real cyclic sub-field X of degree m over F such that $W = (X, K)$.*

For Z is cyclic of degree m over the total real field F, and must contain its complex conjugate field $\bar{Z}$. Hence there is an automorphism T of Z such that $z^{(T)} = \bar{z}$ for every complex number z of Z. Then $z^{(T^2)} = \bar{\bar{z}} = z$ for every z of Z. Thus the automorphism T has order one or two.

If T has order one then $\bar{z} = z$ for every z of Z and $Z = X$ is total real. When T has order 2, then m is even so that n is even, $m = 2n$. But the automorphism group of Z is cyclic and the only automorphism of Z of order two is S^n, S a generating automorphism of Z. Hence $T = S^n$.

I have shown[15] that $Z = F(x)$ such that $Z_n = F(x^2)$ is cyclic of degree n over F, $x^{(T)} = -x$, $x^{(S)} = \beta x$ with β in $F(x^2)$. But then $\bar{x} = -x$, so that Z is a total pure imaginary extension of the total real field Z_n.

We may write $K = F(q)$, $q^2 = \mu < 0$, $\bar{q} = -q$, and define $X = F(x_0)$, $x_0 = qx$. Then $\bar{x}_0 = x_0$ and X is total real. Moreover, X is cyclic of degree m over F with generating automorphism S given by $x_0^{(S)} = qx^{(S)} = q\beta x = \beta x_0$, β in $F(x^2) = F(x_0^2)$ when $x_0^2 = \mu x^2$. This proves Lemma 7 since evidently W is the composite of X and K when X is defined as above.

Since W splits A, it is isomorphic to a cyclic field of degree m in A. Then A is a cyclic algebra and we have proved

THEOREM 3. *Let D be a normal division algebra of degree n over K, a total pure imaginary quadratic extension of a total real algebraic field F, and let $D = A$, $m = n$, or $D = A \times M_2$, $m = 2n$, according as n is odd or even. Then A has a basis*

$$x^i y^j \qquad (i, j = 0, 1, \cdots, m-1) \tag{55}$$

where x satisfies an equation $\phi(x) = 0$ which has degree m, coefficients in F, and the cyclic group for both F and K. The roots of $\phi(x) = 0$ are $\theta^i(x)$ in $F(x)$ and A has the multiplication table

$$y^i x = \theta^i(x) y^i, \qquad y^m = \gamma, \tag{56}$$

with γ in K.

We next prove

THOEREM 4. *Let $n = 2n_0 = 2n_1$, n_0 odd, let D be a normal division algebra of degree n over K, so that $D^n = M \times Q$, where M is a total matric algebra and Q a division algebra of degree two. Suppose that $Q = Q_0 \times K$ where*

$$Q_0 = (1, i, j, ij), \qquad ji = -ij, \qquad i^2 = \alpha, \qquad j^2 = \beta \tag{57}$$

and α, β are total positive quantities of F. Then the conclusion of Theorem 3 holds with $A = D$.

We define $\pi_1, \cdots, \pi_r$ and $p_1, \cdots, p_r$ as before, and let $p_{r+1}, \cdots, p_k$ be any additional prime spots for which Q_0 has p-index 2. Since $F(i)$ is a total real field, the algebra Q_{0p} has index one for every infinite prime spot p of F, and the primes $p_1, \cdots, p_k$ are all finite prime ideals. We may then use Lemma 3 to define Z as before and, by the use of Lemma 7, will have Theorem 4 if we can prove that $W = (Z, K)$ splits $D = A$.

The field Z has p_i-degree n so that its quadratic sub-field Z_2 evidently has p_i-degree two. But then Z_2 splits Q_0 by Lemma 3. Hence the composite W_2 of Z_2 and K is a splitting field of $Q = Q_0 \times K$. By Lemma 3 the π_i-degrees of W_2 are all two for the primes π_i for which Q has π_i-index two. Moreover these primes are precisely the primes π_i for which D has index $2^\nu n_2$, n_2 a divisor of n_0, ν given by (54).

We now consider the π_i-degrees of W. As in the proof of Lemma 6, the π_i-degrees of W are either $n = 2^\nu n_0$ or $2^{\nu-1}n_0$, and these integers are multiples of the π_i-indices of D except possibly where these π_i-indices are $2^\nu n_2$, n_2 a divisor of n_0. But for these latter primes π_i we have proved that W_2 has π_i-degree 2.

The field W has π_i-degree $2^{\nu-1}n_0$ if, and only if, the corresponding field (Z, F_{p_i}) has a quadratic sub-field equivalent to K_{π_i}. But this quadratic sub-field is $(Z_2 \cdot F_{p_i})$ and is equivalent to K_{π_i} if, and only if, the composite of (Z_2, F_{p_i}) and K_{π_i} has degree unity over K_{π_i}. This latter composite is evidently the composite of $W_2 = (Z_2, K)$ and K_{π_i} and has degree two over K_{π_i} by our above assumption on the π_i-degree of W_2.

Hence W has π_i-degree $2^\nu n_0$ at all the prime spots π_i of K for which D has π_i-index $2^\nu n_2$, and has π_i-degree divisible by the π_i-index of D at all the π_i. By Lemma 3 the field W splits D. We have proved Theorem 4.

6. **Riemann matrices of the second kind.** Let ω be a pure Riemann matrix of the second kind, and D its multiplication algebra. Then D is a normal division algebra of degree n over its centrum K. The field K is then known to be an abstract total pure imaginary quadratic extension of a total real field F. Moreover, if ω has genus g then g is divisible by n^2t, t the degree of F over R.

By Theorem 3, the algebra $\mathfrak{A} = D$ or $\mathfrak{A} = D \times M_2$ is a cyclic algebra defined by (55), (56). If $\mathfrak{A} = D \times M_2$ we put

$$\omega_0 = \begin{pmatrix} \omega & 0 \\ 0 & \omega \end{pmatrix} \tag{58}$$

and it is well known that ω_0 is a Riemann matrix with $\mathfrak{A}$ as multiplication algebra. If $\mathfrak{A} = D$ put $\omega_0 = \omega$.

Let $F = R(s)$, $X_0 = R(x, s)$. The field X_0 is cyclic over $F(s)$ and x may always be chosen so that $R(x) = R(x, s)$. Assume that $r = mt$ and that $\xi_1, \cdots, \xi_r$ are the algebraic number roots of the minimum equation of x. The ξ_j are evi-

dently all real, and I have proved[16] that we may then choose a principal matrix C of ω_0 so that x is a *symmetric multiplication* of ω_0. In fact, if x is given by $\nu\omega_0 = \omega_0 N$ then we may take ω_0 of genus g,

$$(59) \qquad \nu = \begin{Vmatrix} \xi_1 I(mg_0) & & & 0 \\ & \xi_2 I(mg_0) & & \\ & & \cdot & \\ & & & \cdot \\ & & & \cdot \\ 0 & & & \xi_r I(mg_0) \end{Vmatrix}, \quad g = mrg_0$$

$$(60) \qquad \omega_0 = \begin{Vmatrix} \tau_1 & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \tau_r \end{Vmatrix} V, \qquad V = \| \xi_j^{k-1} I(2mg_0) \|,$$

$$VNV^{-1} = \begin{Vmatrix} \xi_1 I(2mg_0) & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \xi_r I(2mg_0) \end{Vmatrix},$$

and C such that

$$(61) \qquad VCV' = \begin{Vmatrix} \Gamma(\xi_1) & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \Gamma(\xi_r) \end{Vmatrix}, \quad i\omega_0 C\bar{\omega}_0' = \begin{Vmatrix} i\tau_1 \Gamma(\xi_1)\bar{\tau}_1' & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & i\tau_r \Gamma(\xi_r)\bar{\tau}_r' \end{Vmatrix}$$

so that $CA'C^{-1} = A$.

If b is any multiplication $\beta\omega_0 = \omega_0 B$ of ω_0, then $B^* = CB'C^{-1}$ defines a new multiplication b^* of ω_0 with the properties

$$(b^*)^* = b, \quad (b_1 + b_2)^* = b_1^* + b_2^*, \quad (b_1 b_2)^* = b_2^* b_1^*, \quad \lambda^* = \lambda, \quad q^* = -q$$

for every b, b_1, b_2, of $\mathfrak{A}$ and λ of $R(s)$. In particular, $yx = \theta(x)y$ where $\theta(x)$ is in $R(x)$ and hence $[\theta(x)]^* = \theta(x)$. We compute $(yx)^* = xy^* = y^*\theta(x)$ and $yy^* \cdot \theta(x) = yxy^* = \theta(x) \cdot yy^*$. The quantity $d = yy^*$ is commutative with a generating element $\theta(x)$ of a maximal sub-field $K(x)$ of $\mathfrak{A}$ and, as is well known, is in $K(x)$. Moreover $d^* = d$, so that d is in $R(x)$.

Let y be given by $\eta\omega_0 = \omega_0 Y$. Then $y^* = y^{-1} d(x)$ so that

$$(62) \qquad Y^* \equiv CY'C^{-1} = Y^{-1} d(N),$$

and hence

$$(63) \qquad YCY' = d(N)C.$$

---

[16] Palermo Rendiconti, vol. 55 (1931), pp. 57–115, p. 73.

The matrix η is a non-singular matrix and $i\omega_0 C\bar{\omega}_0'$ is positive definite so that it is

$$\eta(i\omega_0 C\bar{\omega}_0')\bar{\eta}' = i\omega_0(YCY')\bar{\omega}_0' = i\omega_0 d(N)C\bar{\omega}_0' = d(\nu)i\omega_0 C\bar{\omega}_0', \tag{64}$$

where $\bar{Y}' = Y'$ since Y is a rational matrix. By (59)–(61)

$$d(\nu)i\omega_0 C\bar{\omega}_0' = \begin{Vmatrix} d(\xi_1)H_1 & & 0 \\ & \ddots & \\ 0 & & d(\xi_r)H_r \end{Vmatrix} \tag{65}$$

where $H_j = i\tau_j \Gamma(\xi_j)\bar{\tau}_j'$ is a principal minor of $i\omega_0 C\bar{\omega}_0'$ and must be positive definite. Similarly $d(\xi_j)H_j$ must be positive definite. Hence $d(\xi_j) > 0$ and $d(x)$ is a total positive quantity of the total real field $R(x)$.

We have proved that $y^* = y^{-1}d(x)$, $\gamma^* = (y^*)^m = [y^{-1}d(x)]^m = N_{Z/F}(d)\gamma^{-1}$ where $F = R(s)$, $Z = F(x)$. Hence if $\gamma = \gamma_1 + \gamma_2 q$ then

$$\gamma\gamma^* = \gamma_1^2 - \gamma_2^2\mu = N_{Z/F}(d). \tag{66}$$

We have proved

LEMMA 8. *The multiplication algebra of ω_0 is an algebra $\mathfrak{A}$ of Theorem 2.*

We let $W = K(x)$, S be the automorphism of W carrying x to $\theta(x)$, as above, $n = 2n_1$. The field $F(x)$ contains a cyclic sub-field X_4 of degree four over F which is a total real field, and we define W_4 to be the composite of X_4 and K. Write the cyclic algebra $\mathfrak{A} = (W, S, \gamma)$. It is well known that then

$$\mathfrak{A}^{n_1} = M \times (W_4, S, \gamma)$$

where M is a total matric algebra. Now $\mathfrak{A}_4 = (W_4, S, \gamma)$ is a cyclic algebra of degree four over K, and has index two since $\mathfrak{A}$ has index $n = 2n_1$. Moreover if $\gamma = \gamma_1 + \gamma_2 q$ then $\gamma_1^2 - \gamma_2^2\mu = N_{X_4/F}(d_4)$ where $d_4 = N_{Z/X_4}(d)$ is also total positive. Hence $\mathfrak{A}_4$ is an algebra of Theorem 2.

By Theorem 2 there exists a Riemann matrix ω_4 with $\mathfrak{A}_4$ as multiplication algebra. Algebra $\mathfrak{A}_4 = M_2 \times Q$ where Q is a division algebra of degree two over K, and M_2 is a total matric algebra. The Riemann matrix ω_4 is then impure and is in fact necessarily isomorphic to a matrix

$$\begin{pmatrix} \omega_2 & 0 \\ 0 & \omega_2 \end{pmatrix}$$

where ω_2 is a *pure* Riemann matrix with Q as multiplication algebra. It is evident that Q is the algebra of Theorem 4.

I have proved[17] that Q must have the form of *Theorem* 4, and hence that D is a cyclic algebra of that theorem. Our proof of Lemma 8 implies that D is

[17] Loc. cit. (4).

then an algebra of Theorem 2. If n is odd, then this is already implied by Lemma 8. We have proved the

FUNDAMENTAL THEOREM. *The multiplication algebra of any pure Riemann matrix of the second kind is an algebra $\mathfrak{A}$ defined by* (19)–(29). *Conversely every division algebra $\mathfrak{A}$ defined by* (19)–(29) *is the multiplication algebra of a pure Riemann matrix of the second kind.*

7. **The existence of division algebras.** The multiplication algebra D of a Riemann matrix ω of the second kind is a normal division algebra of degree n over its centrum K of degree $2t$. We have proved that D is necessarily an algebra $\mathfrak{A}$, but have not yet shown the *existence of division algebras $\mathfrak{A}$ with arbitrary n and t.*

We let p be any rational prime and $p_1 \neq p$ be a prime of the form $x_1^2 + x_2^2 p$ where x_1 and x_2 are rational integers. Write $\mu = -p$,

$$\Lambda = R(\mu^{\frac{1}{2}}).$$

The integer p_1 does not divide p and hence is not the square of a prime ideal of Λ. But $p_1 = \pi\bar{\pi}$ where $\pi = x_1 + x_2\mu^{\frac{1}{2}} \neq \bar{\pi} = x_1 - x_2\mu^{\frac{1}{2}}$. *Hence the principal ideal (π) is a prime ideal not equal to the principal prime ideal $(\bar{\pi})$.*

By Grunwald's existence theorem (loc. cit.) there exists a cyclic field of degree $2nt$ over R in which (p_1) is a prime ideal.[18] This cyclic field contains a total real sub-field Z of degree n over a total real sub-field F. The field F has degree t over R, Z is cyclic over F, and (p_1) is a prime ideal of Z and F.

We let K be the composite of Λ and F, W be the composite of Z and K. Evidently W is a quadratic extension of Z and the prime ideal p_1 of Z factors in W into $(\pi)\cdot(\bar{\pi})$ where (π) and $(\bar{\pi})$ are prime ideals of W but are not necessarily distinct. If $(\pi) = (\bar{\pi})$, then $\bar{\pi}$ is in (π) so that $\bar{\pi} = a\pi$ where a is an integer of W. But then $\bar{\pi}^{nt} = N_{W/\Lambda}(a)\pi^{nt} = \beta\pi^{nt}$ where β is an integer of Λ. The prime ideal (π) of Λ does not divide $(\bar{\pi})^{nt}$, a contradiction.

We now define the cyclic algebra $\mathfrak{A}$ of (19)–(29) with W as given here, but with

$$\gamma = \pi\bar{\pi}^{n-1}, \qquad d = p_1 > 0,$$

so that $\gamma\bar{\gamma} = (\pi\bar{\pi})^n = N_{Z/F}(p_1)$ and (29) is satisfied. The algebra $\mathfrak{A}$ so defined will be a division algebra if no power of γ less than the n^{th} power is the norm of any quantity of W.

Let then $\gamma^r = N_{W/K}(\beta)$ where β is in W, $0 < r < n$. We may write $\beta = \pi^e\beta_0$ where β_0 is in W and is prime[19] to the prime ideal (π). Then

$$N_{W/K}(\beta) = \pi^{ne}\cdot N_{W/K}(\beta_0) = \pi^{ne}\delta$$

where δ is in K and is prime to π,

$$\gamma^r = \pi^r\bar{\pi}^{r(n-1)} = \pi^{ne}\delta.$$

---

[18] Cf. (14) for this application.

[19] By this we imply that both numerator and denominator of β_0 are prime to (π).

If $e < 0$, then (π) divides δ contrary to our hypothesis that (π) is prime to δ. Hence $e \geqq 0$. But $e \neq 0$ since π divides γ^r. Hence $e > 0$ and $(\pi)^n$ divides $\pi^r \bar{\pi}^{r(n-1)}$. But this is impossible since $r < n$, (π) is a prime ideal and not a divisor of $(\bar{\pi})^{r(n-1)}$.

THEOREM 5. *There exist division algebras* $\mathfrak{A}$ *of* (19)–(29) *with arbitrary* n *and* t.

THE INSTITUTE FOR ADVANCED STUDY
PRINCETON, N. J.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

THE TRANSFORMATIONS EFFECTING THE REDUCTION OF POSITIVE QUATERNARY QUADRATIC FORMS

BY BURTON W. JONES

(Received March 27, 1933)

1. **Introduction.** Two quadratic forms are said to be equivalent if one may be taken into the other by a linear transformation with integral coefficients and of determinant unity, i.e., a unimodular transformation. All forms equivalent to one form are said to form a "class" of equivalent forms. Since two equivalent forms have many properties in common it becomes of prime importance to have some means of selecting a so-called "reduced form" which is representative of the class. Eisenstein's[1] reduction for positive ternary quadratic forms is well known. The author[2] has completed the reduction for positive ternary quadratic forms using Selling's reduction. Though certain general methods of reduction exist[3] there has been found no set of inequalities defining a reduced positive quaternary quadratic form *unique* in its class; that is, such that there is one and only one reduced form in each class. This paper makes accessible the reduction of positive quaternary quadratic forms and develops certain methods that show some promise of applicability to forms of more than four variables.

Before defining more precisely the problem of this paper we shall make for n variables certain definitions and transformations used by Charve[4] for $n = 3$ and 4.

In $f = \sum_{i,\ j=1}^{n} a_{ij} x_i x_j$, $(a_{ij} = a_{ji})$ replace x_i by $x_i - x_{n+1}$ and have

$$\varphi = -\sum_{1=i<j}^{n+1} a_{ij}(x_i - x_j)^2,$$

where

$$a_{i,\ n+1} = -\sum_{j=1}^{n} a_{ij}, \qquad i = 1, 2, \cdots, n+1.$$

It should be noted that f and φ represent the same numbers for integral values of the variables.

---

[1] Dickson, Studies in the Theory of Numbers, pp. 155–180.

[2] Amer. Jour. of Math., vol. 54 (1932), pp. 14–34.

[3] Hermite, Jour. für Math., vol. 40 (1850), p. 281. Minkowski, Jour. für Math., vol. 129 (1905), p. 228.

[4] Annales de l'Éc. Norm. Sup. (2), Vols. 9, 11 (1880, 1882).

Suppose the transformation $x = t'X$, integral and of determinant ± 1 takes f into $F = \sum_{i,\,j=1}^{n} A_{ij} X_i X_j$. Then the substitution of $X_i - X_{n+1}$ for X_i gives

$$\Phi = -\sum_{1=i<j}^{n+1} A_{ij}(X_i - X_j)^2,$$

where

$$A_{i,\,n+1} = -\sum_{j=1}^{n} A_{ij}, \qquad i = 1, 2, \cdots, n+1.$$

Now the transformation $x = tX$ takes φ into Φ if $t_{ij} = t'_{ij}\,(i, j \leqq n)$, $t_{i,\,n+1} = -\sum_{j=1}^{n} t'_{ij}$ $(i = 1, 2, \cdots, n)$ and $t_{n+1,\,i} = 0$ $(i = 1, 2, \cdots, n+1)$.

It is to be noted that subtracting a number from every element of any column of T leaves Φ unaltered.

Throughout this paper, without further remark, T refers to an integral transformation for which the sum of the elements in each row is the same and where $1, -1$ or 0 is the value of the determinant of the matrix T' found by subtracting the bottom element of each column of T from all the elements of that column and then suppressing the last column and row. If the determinant is zero T is said to be "degenerate." We call two T matrices *equal* and, in listing the transformations, consider them as the same matrix, if their T' matrices obtained in the above way are identical. In other words, the T matrices $T_a = (a_{ij})$ and $T_b = (b_{ij})$ are equal if and only if $a_{ij} - b_{ij} = r_j$ for $i, j = 1, \cdots, n+1$, where r_j is an integer (positive, negative or zero) depending on j but not on i. Furthermore, we shall not distinguish between a matrix T and $-T$. It may quickly be established that $(T_1\,T_2)' \equiv T_1'T_2'$ and that $T_a = T_c$ and $T_b = T_d$ in the above sense implies that $(T_aT_b)' \equiv (T_cT_d)'$. It is also interesting to note that the class of T matrices is closed under addition and multiplication as well as the operation of adding or subtracting integers from columns. Finally note that with each transformation T' of determinant ± 1 are associated numerous matrices T and with each T is associated one T'.

Let $s = (\tfrac{1}{2}) \sum_{i=1}^{n+1} a_{ii} = -\sum_{i<j}^{n+1} a_{ij} > 0$, where the second equality follows from the definition of the a's. Under the transformation T we have $x_i - x_j = \sum_{k=1}^{n+1} (t_{ik} - t_{jk})X_k$ and the coefficient of X_k^2 in Φ is $-\sum_{i<j}^{n+1} a_{ij}\,(t_{ik} - t_{jk})^2$. Thus

$$\Delta s = s' - s = -\Sigma c_{ij}a_{ij},$$

where $2c_{ij} = \sum_{k=1}^{n+1} (t_{ik} - t_{jk})^2 - 2$ and $i, j = 1, 2, \cdots, n+1$ with $i < j$.

Define $c_{ji} = c_{ij}$ but leave c_{ii} undefined. We regard c_{ij} and c_{ji} as two names for the same thing.

Now the last sum is always even and thus $c_{ij} < 0$ only if the two rows are identical and the matrix therefore degenerate; $c_{ij} = 0$ if and only if the differences between corresponding elements of the i^{th} and j^{th} rows are $1, -1, 0, 0, \cdots, 0$ in some order.

Selling[5] for three variables and Charve[4] for three and four called a form f *reduced* if s is a minimum, that is, if for every transformation T (non-degenerate),

$$\Delta s = -\sum_{i<j}^{n+1} c_{ij}a_{ij} \geqq 0.$$

Since it seems preferable to consider a reduced form to be equivalent to no other reduced form, we shall call a form f *semi-reduced* when s is a minimum.

In this connection three problems arise: first, what are the conditions on the coefficients of a form f so that it shall be semi-reduced; second, for any given semi-reduced form, what transformations T leave s unaltered; third, given a class of equivalent semi-reduced forms, what further restrictions may be put on the coefficients in order to define a *unique* form representative of the class. We shall call this unique form *reduced*. *Every form is thus equivalent to some reduced form and no two reduced forms are equivalent.*

The first of the three problems above was considered for quaternaries by Charve but though the results were correct (if given the proper interpretation) his proofs are at times sketchy.[6] This is here put on a sound basis and the second problem is completely solved for quaternaries. In fact, every transformation T leaving the s of a semi-reduced form unaltered is shown to be one of four types multiplied on the left by an automorph unless just one $a_{ij} > 0$ in which case there are seven types. From the theory of this paper one can immediately write down the transformations T leaving unaltered the s of any given semi-reduced form f. In order to write down these transformations it is necessary merely to know what a's vanish or satisfy the conditions in Fundamental Theorems 3 and 2. These transformations include necessarily all the automorphs of f. The complete reduction for quarternaries should follow from this and will be developed in a later paper.

Some of the parts of the reduction are easily extended to n variables but others become increasingly complex when applied to more than four variables.

2. We shall find it useful to obtain an expansion for d, the symmetric determinant $|a_{ij}|$ $(i, j = 1, 2, \cdots, n)$ in terms of a_{ij} $(i < j; i, j = 1, 2, \cdots, n+1)$.

[5] Jour. de Math. (3), Vol. 3 (1877).

[6] After this paper was submitted for publication a paper entitled "Verallgemeinerung der Sellingschen Reduktionstheorie" by N. Hofreiter appeared in *Monatshefte für Mathemaitk and Physik*, vol. 40 (1933), pp. 393–406, in which Charve's errors were corrected. But Hofreiter was *concerned only with the form reduced in the Selling sense.*

Thus the second form of the determinant is the same as the first except for the elements a_{ii} along the principal diagonal, these being replaced, according to earlier definitions, by $-\Sigma a_{ij}$ the sum being over $j = 1, 2, \cdots, n+1$ with $j \neq i$. Since $a_{ij} = a_{ji}$ we note that any element $a_{ij} (i \neq j)$ in the second form of the determinant occurs by itself in the j^{th} row and i^{th} column as well as in the i^{th} row and j^{th} column and, as a member of a sum, in the principal diagonal in the i^{th} and j^{th} rows. In the expansion of d in terms of the second set of a's consider the coefficient of $r = a_{12}a_{23}a_{34} \cdots a_{k1}$ where $k < n+1$. All the terms in these a's occur in the first k rows and columns of d which we shall call d_k. Using Laplace's expansion we see that the coefficient of r in d is the complementary minor of d_k multiplied by the coefficient of r in d_k. We prove that this latter coefficient is zero. Suppose a_{12} comes from the second row, first column and a_{23} from the third row, third column. Then a_{34} comes from the fourth row, fourth column, etc., a_{k1} from the first row and first column which cannot be. Similar reasoning shows that every term r is obtained either by multiplying elements of the principal diagonal or elements none of which are in the principal diagonal. The coefficient of this term in the principal diagonal is $2(-1)^k$, for if a_{12} comes from the first row, a_{k1} must come from the k^{th}, a_{jk} from the j^{th} and so forth, while if a_{12} comes from the second row, a_{23} must come from the third and so forth. Consider next the terms $\Sigma(-1)^i a_{12}a_{23} \cdots a_{k1}$, products no element of which is taken from the principal diagonal, i being the total number of inversions in rows and columns. Reasoning as above we see that the sum contains two terms for each of which $i = k - 1$. Thus the coefficient of r in d is zero. We call r a *k-cycle*; that is, any product of k elements which may be written with subscripts in cyclic order. From reasoning similar to that above it may be seen that there is no k-cycle involving a subscript $n+1$. Thus, in the expansion of d all terms containing a k-cycle vanish, $0 < k < n+1$.

We prove that any product of a_{ij}'s $(i < j)$ every a of which is not taken from the principal diagonal has, in the expansion of d, the coefficient zero. Consider the term $\tau = a'_{i_1 j_1} \cdot a'_{i_2 j_2} \cdots a'_{i_n j_n}$ where a'_{ij} is the term in the i^{th} row and j^{th} column of d. Suppose in $n - k$ of these $i = j$ and $\tau = a'_{i_1 j_1} \cdot a'_{i_2 j_2} \cdots a'_{i_k j_k} P$, where P is the product of elements in the principal diagonal and $i_1, \cdots, i_k$ and $j_1, \cdots, j_k$ are each permutations of a certain set of k numbers and therefore containing a t-cycle where $t \leqq k$. Thus the coefficient of this term in the expansion of d is zero.

Finally we prove $d = (-1)^n \Sigma a_{1j_1} a_{2j_2} \cdots a_{nj_n}$ where the summation extends over all sets of values of j_1 to j_n taken from the numbers $1, 2, \cdots, n+1$ (some may be repeated) no term of which contains a k-cycle, $k \leqq n$. From the above it follows that $d = (-1)^n \prod_{i=1}^{n} (a_{i1} + a_{i2} + \cdots + a_{in+1} - a_{ii}) - N$ where N is the sum of all the terms containing a cycle. Reordering the subscripts so that the first indicates the row we see that all such terms as indicated at the beginning of the paragraph occur. Consider the term $r = a_{1j_1} a_{2j_2} \cdots a_{nj_n}$, in the

expansion of d above. To complete the proof we show that unless every first subscript indicates the row from which the a is taken, the term contains a cycle. Suppose $a_{j_1 1}$ came from the j_1^{th} row; then one of the other elements must come from the first row and thus $1 =$ one of $j_2, j_3, \cdots, j_n$. Permute subscripts if necessary and take $j_2 = 1$. Then $2 =$ one of $j_3, \cdots, j_n$. Call it j_3; so continuing we find that $j_1, j_2, \cdots, j_n$ is some permutation of $1, 2, \cdots, n$ and the term contains a cycle. In other words, the coefficient of any r in d is $(-1)^n$ unless it contains a cycle.

COROLLARY 1. *$d = 0$ if $n(n-1)/2 + 1$ of the a_{ij} $(i < j)$ are zero, for the number of these a's is $n(n+1)/2$ and every product of n distinct a's will vanish if $n(n+1)/2 - n + 1$ vanish. Also d will vanish if all products of n a's vanish except those containing a cycle.*

COROLLARY 2. *From the form of the determinant it is apparent that if for any i, $a_{ij} = 0$ for every $j \neq i$, then $d = 0$. Note that $a_{ii} \leqq 0$ for any i implies that f is not positive.*

3. **Definitions.** Define

$$\begin{bmatrix} -1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = I_{-1,5}$$

and similarly $I_{-i,j}$ as the matrix formed by subtracting 1 from each element of the j^{th} column of the identity matrix and then changing the signs of all the elements of the i^{th} row. $I_{-i,j}$ is a matrix of type T and it may be seen that if any matrix T, every element of whose j^{th} row vanishes, is multiplied on the left by $I_{-i,j}$, the resulting matrix is T with the signs of the elements of the i^{th} row changed. Δs for $I_{-1,5}$ is $-2(a_{12} + a_{13} + a_{14})$.

Define similarly $I_{-ij,k}$ and note a similar effect. For $I_{-ij,k}$, Δs is $-2(a_{ir} + a_{is} + a_{jr} + a_{js})$ where i, j, k, r, s is a permutation of 1, 2, 3, 4, 5. Transformations $I_{-i,j}$ or $I_{-ij,k}$ are called transformations of the type I_-.

Define I_p, where p is any substitution group on five or less numbers, to be that matrix obtained by permuting the rows of the identity matrix according to the plan p. For instance,

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} = I_{123,45}.$$

Multiplication of any matrix T by I_p on the left permutes the rows of T by p and multiplication on the right permutes the columns of T by p^{-1}.

A transformation T is said to be *permissible* for f if the a_{ij} $(i \neq j)$ of φ are

such that the Δs of T is zero. A transformation T_1 is said to be permissible for T if the Δs of T_1 contains no a_{ij} not in the Δs of T.

4. **Various lemmas.**

LEMMA 1. *If the first three rows of a matrix T are* $1\ -1\ 0\ 0\ 0$, $-1\ 0\ 1\ 0\ 0$ *and* $t_{31}t_{32}t_{33}t_{34}t_{35}$, *if the fourth or fifth row is zero and if* $c_{13} = c_{23} = 0$, *the matrix is degenerate.*

PROOF: Since $(t_{21} - t_{31})^2 \leqq 1 \geqq (t_{11} - t_{31})^2$, $t_{31} = 0$. If $t_{32} = 0$, then $c_{13} = 0$ implies $t_{33} = t_{34} = t_{35} = 0$. Thus the third row is $0\ -1\ 1\ 0\ 0$, it is a linear combination of the other two and the matrix is degenerate.

LEMMA 2. *If the first two rows of a non-degenerate transformation T are* $1\ -1\ 0\ 0\ 0$ *and* $t_{21}t_{22}t_{23}t_{24}t_{25}$, *if* $c_{12} = 0$ *and if one of the last three rows is zero, then, after multiplying by* $I_{-2,1}$ *(if necessary) we can make the second row one of* $0\ -1\ a\ b\ c$, $1\ 0\ -a\ -b\ -c$, $1\ -1\ a\ b\ c$, where *in the first two cases* a, b, c *are* $1, 0, 0$ *in some order and in the third* $1, -1, 0$ *in some order, where the signs of these three elements may be changed at will, e.g., by multiplication by* $I_{-2,1}$ *one may change* $1\ -1\ 1\ -1\ 0$ *to* $1\ -1\ -1\ 1\ 0$. *Multiplication by* $I_{-23,1}$ *will have the same effect.*

PROOF: Subtract 1 from the elements of the first column of T and add 1 to the elements of the second. The new second row is $t_{21} - 1, t_{22} + 1, t_{23}, t_{24}, t_{25}$ and, by changing the sign of the second row and permuting the last three elements if necessary we can make it one of $-1\ 1\ 0\ 0\ 0$, $-1\ 0\ 1\ 0\ 0$, $0\ 1\ -1\ 0\ 0$, $0\ 0\ 1\ -1\ 0$ yielding the three possibilities mentioned above, since the t_{2i} are not all zero.

LEMMA 3. *Given a non-degenerate transformation T having the following properties*

1. $1 < c_{12} \geqq c_{ij}$ *for all* i, j.
2. $c_{ij}c_{ik}c_{il}c_{im} = 0$ *for all* i, j, k, l, m.
3. *At least four c's vanish.*
4. *If just four c's vanish, three of the four are not of the form* c_{ij}, c_{jk}, c_{ki}.

Then there exist I_-'s such that $\cdots I'''_-I''_-I'_-T = T'$ *where I'_- is permissible for T, I''_- for I'_-T, etc., and $\Delta s'$ (the Δs of T') either contains one less c_{ij} than the Δs of T or every c_{ij} of $\Delta s'$ is $\leqq 1$. Throughout the statement of this lemma i, j, k, l, m is a permutation of* 1, 2, 3, 4, 5.

Note that if, for instance, $I_{-1,5}$ is permissible for a transformation T, the only c's in $I_{-1,5}T$ different from the corresponding c's in T are c_{12}, c_{13}, c_{14} which are all contained in the Δs of T due to the fact that $I_{-1,5}$ is permissible for T. Thus, if any I_- is permissible for a transformation T, the Δs of I_-T has no c's in it which did not occur in the Δs of T. It follows that if a c occurs in the Δs of T which does not occur in the Δs of I_-T, the latter contains at least one less c than the former.

First, if $c_{13} = c_{23} = 0$ and the matrix is non-degenerate there is an I_- for which $c'_{12} = 0$ where c'_{12} is the c_{12} of the matrix I_-T. If $c_{1j} = c_{2j} = 0$ with $j = 4$ or 5, make the j^{th} row zero and see by lemma 1 that the matrix is degenerate. Otherwise write T with the third row zero and take as the first two rows

$1\ -1\ 0\ 0\ 0$ and $0\ 1\ -1\ 0\ 0$. If $c_{i4}c_{i5} \neq 0$ where $i = 1$ or 2 use $I_{-i,3}$ to make $c'_{12} = 0$. If $c_{14}c_{15} = 0 = c_{24}c_{25}$ take $c_{14} = 0 = c_{25} \neq c_{15}c_{24}$ and use $I_{-14,3}$ to make $c'_{12} = 0$ unless $c_{45} = 0$. There remains the typical case $c_{14} = 0 = c_{25} = c_{45} \neq c_{15}c_{24}$. Write the first three rows of T as $0\ 0\ 0\ 0\ 0$, $-1\ 2\ -1\ 0\ 0$, $-1\ 1\ 0\ 0\ 0$ and using hypothesis 1 of the lemma see that there are the following possibilities for the fourth row: $-1\ 0\ 0\ 1\ 0$, $0\ 1\ 0\ -1\ 0$, $0\ 0\ -1\ 1\ 0$. Since $c_{25} = 0$, $t_{52} \geqq 1$. Since T is not degenerate, $t_{55} \neq 0$. Thus, since $c_{45} = 0$, the fourth row is $0\ 1\ 0\ -1\ 0$. If $t_{52} = 2$ the fifth row is $0\ 2\ 0\ -1\ -1$ and $c_{25} \neq 0$. If $t_{52} = 1$, the fifth row is $-1\ 1\ 0\ -1\ 1$ or $0\ 1\ -1\ -1\ 1$ and $c_{25} \neq 0$.

Second: $c_{13} = c_{14} = c_{25} = c_{35} = 0 \neq c_{15}c_{23}c_{24}$. If $c_{23} > 1$, $c_{25} = c_{35} = 0$ implies $c_{23} = 2$ and multiplication by $I_{-2,5}$ makes $c'_{23} = 0$. Thus $c_{23} = 1$ remains and the second, third and fifth rows may be written $1\ -1\ 1\ -1\ 0$, $0\ 0\ 0\ 0\ 0$, $1\ -1\ 0\ 0\ 0$. $c_{13} = c_{35} = 0$ implies $c_{15} = 1$ or 2.

a) If $c_{15} = 1$ we can use $I_{-2,5}$ if necessary to make $c_{12} = 1$ and have as the first row $0\ 0\ 1\ 0\ -1$ and then the only possible c's $\geqq 1$ are c_{24}, c_{45} and c_{34}. If $c_{34}c_{45} \neq 0$, use lemma 2 with $I_{-4,1}$ and see that the fourth row may be taken to be $0\ 1\ 1\ -1\ -1$ or $1\ 0\ 1\ -1\ -1$; in the latter case all $c_{ij} \leqq 1$ while in the former the multiplication by $I_{-14,3}$ and $I_{-2,5}$ in succession makes all $c_{ij} \leqq 1$. If $c_{45} = 0$ and $c_{34} \geqq c_{24}$ the hypotheses of the theorem with subscripts 1 and 2 replaced by 3 and 4 are satisfied and the first case above with c_{23} replaced by c_{14} supplies the proof. If $c_{45} = 0$ and $c_{34} < c_{24}$ we have a similar proof since $c_{25} = c_{45} = 0$. If $c_{34} = 0 \neq c_{45}$ use $I_{-14,3}$ to make $c'_{45} = 0$ unless the matrix is degenerate.

b) If $c_{15} = 2$, lemma 1 shows that the matrix is degenerate unless $c_{45} \neq 0$. If $c_{45} \neq 0$ use $I_{-14,3}$ to make $c'_{15} = 0$.

Third: $c_{13} = c_{24} = 0 \neq c_{14}c_{15}c_{23}c_{25}$ and two of $c_{34}c_{35}c_{45}$ vanish. From symmetry we take $c_{35} = 0$ and write the first, third and fifth rows as follows: $1\ -1\ 0\ 0\ 0$, $0\ 0\ 0\ 0\ 0$, $0\ 0\ 1\ -1\ 0$. If $c_{34} = 0 \neq c_{45}$, $I_{-24,3}$ can make c_{14} or c_{45} vanish. If $c_{34} = 0 = c_{45}$, $I_{-1,3}$ can make c_{14} vanish unless the fourth row is $0\ 0\ 1\ 0\ -1$; write the matrix with the fourth row zero and see that $I_{-2,4}$ makes c_{23} or c_{25} vanish unless the matrix is degenerate. If $c_{45} = 0 \neq c_{34}$ write the matrix with the fifth row zero and see that $I_{-24,5}$ makes c_{34} vanish unless $t_{43} = t_{44} = 0$. The non-degeneracy of the matrix implies $t_{41}t_{42} = 0$ and we may take as the fourth row $1\ 0\ 0\ 0\ -1$. Then write T with the fourth row zero and see that there is trouble only when the new second row is $0\ -1\ 1\ 0\ 0$. In this case multiplication by $I_{-13,5}$ and then $I_{-1,3}$ reduces T to one in which every $c_{ij} \leqq 1$.

Finally we show that the above include all the typical cases. If $c_{1j} = c_{2j} = 0$ for any $j = 3$, 4 or 5 the first case disposes of it. If $c_{i3} = c_{i4} = c_{i5} = 0$ where $i = 1$ or 2 then by symmetry take $i = 1$ and, since hypothesis 2 implies that some c with a subscript 2 vanishes, this reduces to the first case. If $c_{i3} = c_{i4} = c_{i5} = 0$ is not true for $i = 1$ or $i = 2$ but $c_{ij} = c_{ik} = 0$ for some $i = 1$ or 2 where j and k are two of 3, 4, 5, from symmetry take $c_{13} = c_{14} = 0 \neq c_{15}$. If this is not to be included in the first case we have also $c_{25} = 0 \neq c_{23}c_{24}$. The second case covers this possibility when $c_{35} = 0$. Treatment for $c_{45} = 0$ is the same. If

however $c_{35}\,c_{45} \neq 0$, the only c's which vanish are $c_{13}\,c_{14}\,c_{25}$ and possibly c_{34} which is ruled out by hypothesis 4. Lastly, if $c_{ij} = c_{ik} = 0$ for no $i = 1$ or 2 and j, k two of 3, 4, 5 we may from symmetry take $c_{13} = c_{24} = 0 \neq c_{14}\,c_{15}\,c_{23}\,c_{25}$ and hypothesis 3 implies that at least two of c_{34}, c_{35}, c_{45} must vanish. This is covered by the third case.

LEMMA 4. *Any non-degenerate transformation T satisfying hypotheses* 2, 3 *and* 4 *of lemma* 3 *can be reduced by a multiplication on the left by permissible $I_-'s$ (as in lemma* 3*) to a transformation in which every $c_{ij} \leqq 1$.*

From lemma 3, if any $c_{ij} > 1$, there are transformations I_- yielding in the manner prescribed a transformation T' where $\Delta s'$ contains no non-vanishing c_{ij} which did not occur in the Δs of T, and either every $c'_{ij} \leqq 1$ or $\Delta s'$ contains one less c_{ij}. Since Δs contains no more than six non-vanishing c's, after a few repetitions of the process we have a matrix T' having every $c_{ij} \leqq 1$.

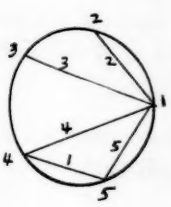
LEMMA 5. *A non-degenerate transformation T satisfying hypotheses* 2, 3 *and* 4 *of lemma* 3 *and whose every $c_{ij} \leqq 1$, is either equal to $\pm I_p$ or equal to plus or minus a matrix T whose every row is $-1\ -1\ 0\ 0\ 0$ in some order.*

If $c_{ij} \leqq 1$ a subtraction of a largest element of each column from every element of that column yields a matrix whose terms are all 0 or -1. Since, in any matrix T, $\sum_{j=1}^{5} (t_{ij} - t_{kj}) = 0$ for every i and k, the number of -1's in each row must be the same. If less than two elements of each row are zero, 1 may be added to every element of the matrix, the matrix multiplied by -1 and a matrix obtained in which three or more elements of every row are zero and the others -1.

LEMMA 6. *Every non-degenerate matrix whose every row is $-1\ -1\ 0\ 0\ 0$ in some order may be obtained by permuting the rows and the columns of the four transformations A, B, C, D below. Note that this method is applicable for n variables.*

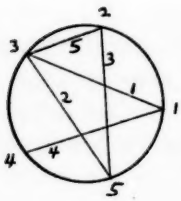
A picture of each such transformation may be made as follows. Draw a circle with the points 1, 2, 3, 4, 5 marked on it. If the first row of T is $-1\ -1\ 0\ 0\ 0$ draw a line from point 1 to point 2 and label it 1 and so proceed. A permutation of the rows (columns) of a matrix permutes the numbers on the lines (points) of the figure and vice versa. We will thus have five lines connecting five points, each point being the end point of at least one line since the matrix is non-degenerate. No point can be an end-point of more than four lines. The possible distributions of ten end-points among five points are 4 2 2 1 1, 3 3 2 1 1, 3 2 2 2 1, 2 2 2 2 2.

We have the following typical figures and transformations.



$$\begin{bmatrix} 0 & 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 & -1 \end{bmatrix} = A_{312};\ -\Delta s = a_{31} + a_{12}$$

$$\begin{bmatrix} -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & -1 & 0 \end{bmatrix} = B_{3124};\ -\Delta s = a_{31} + a_{12} + a_{24}.$$



$$\begin{bmatrix} -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & -1 & 0 & 0 & -1 \\ -1 & 0 & 0 & -1 & 0 \\ 0 & -1 & -1 & 0 & 0 \end{bmatrix} = C_{13452};\ -\Delta s = a_{13} + a_{34} + a_{45} + a_{42}.$$

$$\begin{bmatrix} -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & -1 & 0 & -1 & 0 \\ -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 \end{bmatrix} = D_{12453};\ -\Delta s = a_{12} + a_{24} + a_{45} + a_{53} + a_{31}.$$

The other diagram for the distribution 3 2 2 2 1 yields a degenerate matrix.

Define A_{ijk} to be that matrix obtained from A_{312} by permuting the rows *and* columns according to the plan

$$\begin{pmatrix} 3 & 1 & 2 \\ i & j & k \end{pmatrix}.$$

Proceed similarly for B, C and D.

It is easily verified that the only permutation of rows which leaves $-\Delta s$ for A_{312} unaltered is (32) and $A_{312} = A_{213}$; $-\Delta s$ for B_{3124} is left unaltered only by the permutation of rows (34)(12) and $B_{3124} = B_{4213}$; $-\Delta s$ for $C_{134,\,52}$ is left unaltered only by the permutation (52) of rows and $C_{134,\,52} = C_{134,\,25}$. However the Δs for D_{12453} is left unaltered by every permutation of rows which leaves the cycle (12453) unaltered. Each such permutation of rows and the corresponding columns alters D according to the following laws: $D_{ijklm} = D_{jklmi}\, I_{mkilj}$ and $D_{mlkji} = D_{jklmi}$.

5. Viewing transformations A and B we see

THEOREM 1. *For any positive quaternary quadratic form f there is a transformation T decreasing $-\Sigma a_{ij}$ (the sum over $i, j = 1, 2, 3, 4, 5$; $i < j$) unless the following inequalities hold:*

(1) $$a_{ij} + a_{ik} \leqq 0,$$

(2) $$a_{ji} + a_{ik} + a_{kl} \leqq 0 \text{ for every } i, j, k, l \text{ four of } 1, 2, 3, 4, 5.$$

THEOREM 2. *If a positive quadratic quaternary form f has coefficients satisfying* (1) *and* (2), $I_{-i,j}$ *or* $I_{-ij,k}$ *is an automorph if it is permissible for* f.

PROOF: Suppose $I_{-1,5}$ is permissible, that is, $a_{12} + a_{13} + a_{14} = 0$. This implies $(a_{12} + a_{13}) + (a_{12} + a_{14}) + (a_{13} + a_{14}) = 0$. From (1) therefore, $a_{12} + a_{13} = a_{12} + a_{14} = a_{13} + a_{14} = 0$ and $a_{12} = a_{13} = a_{14} = 0$. Thus $f = a_{15}(x_1 - x_5)^2 + \Sigma a_{ij}(x_i - x_j)^2$ is left unaltered by $I_{-1,5}$, the sum being over all i and j for which $5 \geqq j > i \geqq 2$.

Suppose $I_{-12,5}$ is permissible. Then $a_{13} + a_{14} + a_{23} + a_{24} = 0$ and from (1) $a_{13} + a_{14} = a_{13} + a_{23} = a_{14} + a_{24} = a_{23} + a_{24} = 0$ while (2) implies $a_{13} + a_{23} + a_{24} = 0 = a_{14}$ and thus $a_{13} = a_{14} = a_{23} = a_{24} = 0$ and f is left unaltered.

FIRST FUNDAMENTAL THEOREM: *If in the φ of a positive quaternary quadratic form f every $a_{ij} \leqq 0$, the form f is semi-reduced and the non-degenerate transformations T leaving s unaltered in value are all and only the transformations obtained by permuting the columns of the following matrices: permissible (for f) transformations of the types A, B, C, D (that is, whose Δs's contain only those a's of f which vanish) multiplied on the left by I and by transformations of the types I_- permissible for f.*

PROOF: It follows from the first paragraph of the proof of lemma 3 that all the transformations I_- permissible for the transformations at the various stages must be permissible for f if $'T$ is. Since for every transformations T, $c_{ij} \geqq 0$ its Δs cannot be negative if every $a_{ij} \leqq 0$. Thus the form is semi-reduced. For the Δs of a transformation T to vanish every a_{ij} which it contains must vanish and thus T must satisfy the conditions imposed in the second, third and fourth hypotheses of lemma 3. Lemmas 4, 5 and 6 then apply to prove the theorem.

COROLLARY: *Aside from a reordering of the variables all forms equivalent to a semi-reduced form f and whose s is equal to the s of f may be found by applying to f permissible transformations of the types A, B, C and D.*

EXAMPLE: Suppose $a_{12} = a_{15} = a_{23} = a_{24} = 0$ and all the other a_{ij} $(i \neq j)$ are negative. The transformations are $A_{215}, A_{123}, A_{124}, A_{324}, B_{5123}, B_{5124}, C_{512,34}$, $I_{-2,5}$ and the product of each on the left by $I_{-2,5}$. It may be noted that diagrams are of some assistance in this connection.

6. **We consider the case: $a_{12} > 0$ and (1) and (2) true.** Note

$$\Delta s = (-c_{12} + c_{13} + c_{14} + c_{15} + c_{23} + c_{24} + c_{25})a_{12} - \Sigma\, c_{ij}(a_{12} + a_{ij})$$
$$- c_{35}a_{35} - c_{45}a_{45} - c_{34}a_{34}, \text{ where } i = 1, 2; j = 3, 4, 5.$$

$$\Delta s = (-c_{12} + c_{13} + c_{14} + c_{15} + c_{23} + c_{24} + c_{25})a_{12}$$
$$+ (-c_{34} + c_{13} + c_{14} + c_{35} + c_{32} + c_{42} + c_{45})a_{34} - \Sigma\, c_{ij}\,(a_{12} + a_{34} + a_{ij})$$
$$- \Sigma\, c_{i5}(a_{12} + a_{i5}) - \Sigma\, c_{j5}(a_{34} + a_{j5}), \text{ where } i = 1, 2; j = 3, 4.$$

In the first equation, if $a_{12} > 0$ and $a_{34}, a_{35}, a_{45} \leqq 0$, (1) implies that all terms but the first are positive or zero and therefore that Δs is positive if the coeffi-

cient of a_{12} is positive. Similarly in the second equation if a_{12}, $a_{34} > 0$, (1) and (2) imply that Δs is positive if the coefficient of a_{12} and of a_{34} is positive. Thus we seek the transformations T for which

$$c_{12} \geqq c_{13} + c_{14} + c_{15} + c_{23} + c_{24} + c_{25}. \tag{3}$$

Throughout the section up to the second fundamental theorem $i = 1$ or 2; $j = 3, 4$ or 5; $k = 1, 2, 3, 4$ or 5. (3) holds if and only if

$$\sum_{k=1}^{5} (t_{1k} - t_{2k})^2 \geqq \sum_{i,j,k} (t_{ik} - t_{jk})^2 - 10.$$

Write $\psi(k) = -(t_{1k} - t_{2k})^2 + \sum_{i,j} (t_{ik} - t_{jk})^2 - 2.$

We prove that no non-degenerate matrix T contains a column k for which $\psi(k) < 0$ and find for what columns $\psi(k) = 0$. Fix k, write $d_{ij} = t_{ik} - t_{jk}$, note

$$t_{1k} - t_{2k} = d_{1j} - d_{2j} \tag{4}$$

and have

$$\begin{aligned}\psi(k) &= \sum_j \{-(d_{1j} - d_{2j})^2 + 3d_{1j}^2 + 3d_{2j}^2\}/3 - 2 \\ &= \sum_j (d_{1j} + d_{2j})^2/3 + \sum_j (d_{1j}^2 + d_{2j}^2)/3 - 2.\end{aligned}$$

First if $\psi \leqq 0$, no d is greater than 1. For, if two d's were greater than 1, the second sum is greater than 2; while if $d_{13}^2 > 1$ and $d_{23}^2 \leqq 1$, for example, the first sum would, from (4), be $\geqq 1$.

Second, if $t_{1k} = t_{2k}$, from (4), $d_{1j} = d_{2j}$. Then at least four d's vanish if $\psi \leqq 0$. If all the d's vanish, all elements in the k^{th} column are equal and the matrix is degenerate. If all but two vanish, $\psi = 0$ and by subtracting t_{1k} from each element of the column and multiplying by -1 if necessary we have a column U_j whose element in the j^{th} row is 1 and whose other elements are all zero.

Third, if $t_{1k} - t_{2k} = \pm 1$, just three d's vanish and $\psi = 0$. If three d's with a subscript in common vanish, we have as above U_i. Otherwise we have U_{ij} whose elements in the i^{th} and j^{th} columns are 1 and whose other elements vanish.

Finally, if $t_{1k} - t_{2k} = \pm 2$ then $d_{1j} = \pm 1$, $d_{2j} = \mp 1$ and we have U_{12}: $\pm 1 \mp 1\ 0\ 0\ 0$.

Thus, the inequality in (3) holds for no transformation T and the equality sign holds only when T can be put into a form such that each column is one of U_i, U_j, U_{ij}, U_{12} or its negative. From lemmas 5 and 6 every transformation whose every column is $\pm U_i$, $\pm U_j$ or $\pm U_{ij}$ is of the form A, B, C or D.

Note that, since $a_{12} > 0$ and (1) implies that $a_{ij} < 0$ and see that of the transformations A, B, C, D we need consider in this connection only those

whose Δs contains a_{12} and those whose Δs contains no a with a subscript 1 or 2 or contains some positive a.

It remains to consider the matrices T containing a column U_{12}. Note, from the proof of lemma 5, that the number of non-vanishing elements in each of the five rows is of the same parity and see that, aside from the order, the possible combinations of the U's are $U_{12}U_iU_{ij}U_mU_l$, $U_{12}U_iU_{ij}U_{pm}U_{pl}$, $U_{12}U_{1j}U_{2l}U_jU_l$, $U_{12}U_iU_{ij}U_{im}U_{il}$ where i, p is a permutation of 1, 2 and l, m, j is a permutation of 3, 4, 5. Giving the subscripts particular values, fixing the sign of one U arbitrarily and determining the signs of the others so that the sum of the elements in each row is the same (that is, so that when one row is made to vanish the sum of the elements in each of the others is zero) we have the following typical transformations:

$$U_1U_{12}U_{13}U_4U_5 = I_{-1,3}A_{415},$$

$$U_1U_{12}U_{24}U_{23}U_{15} = I_{-1,5}C_{521,34},$$

$$U_{13}U_{24}U_3U_4U_{12} = I_{-24,5}B_{1432},$$

$$U_1U_{12}U_{13}U_{14}U_{15} = \begin{bmatrix} 1 & 1 & -1 & -1 & -1 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} = E_{12,345}$$

For $E_{12,345}$, $-\Delta s = 3a_{12} + a_{13} + a_{14} + a_{15}$. Now $\Delta s = 0$ when (1) holds only if $a_{12} = -a_{13} = -a_{14} = -a_{15}$. In this case trial shows that E is an automorph.

Finding Δs for each of the transformations above we have

THE SECOND FUNDAMENTAL THEOREM: *If* $a_{ij} > 0$, *if* (1) *and* (2) *hold and if* a_{pl}, a_{pm} *and* $a_{lm} \leqq 0$, *a quaternary quadratic form* f *is semi-reduced*[7] *and the transformations leaving* s *unaltered are obtained by permuting the columns of the following where* i, j, p, l, m *is a permutation of* 1, 2, 3, 4, 5:

Condition	Transformation
1. $a_{pl} = a_{pm} = 0$	A_{lpm}
2. $a_{ij} = -a_{ip}$	A_{jip}
3. $a_{ij} + a_{ip} = 0 = a_{pl}$	B_{jipl}
4. $a_{ij} + a_{jm} = 0 = a_{mp} = a_{ml}$	$C_{ijm,\,pl}$
5. $a_{ij} = -a_{ip} = -a_{im}$	$I_{-i,\,l}A_{mip}$
6. $a_{ij} = -a_{ip} = -a_{im} = -a_{jl}$	$I_{-i,\,l}C_{lji,\,pm}$
7. $a_{ij} = -a_{ip} = -a_{jl}$; $a_{pl} = 0$	$I_{-jp,\,m}B_{iplj}$
8. $-a_{ij} = a_{ip} = a_{il} = a_{im}$	$E_{ij,\,plm}$ (automorph).

[7] By a proof which is not long, f may be shown under these conditions to be positive if its determinant is not zero.

THE THIRD FUNDAMENTAL THEOREM: *If* a_{ij}, $a_{lm} > 0$ *and* (1) *and* (2) *hold, a quaternary quadratic form* f[8] *is semi-reduced and the transformations leaving* s *unaltered are permutations of the columns of the following where* i, j, l, m, p *is a permutation of* 1, 2, 3, 4, 5:

Condition	Transformation
1. $a_{ij} + a_{ip} = 0$	A_{jip}
2. $a_{lm} + a_{lp} = 0$	A_{mlp}
3. $a_{ij} + a_{lm} + a_{il} = 0$	B_{jilm}

5, 6 and 8 above do not occur here because one of a_{il}, a_{jl}, a_{im}, a_{jm} is equal to $-a_{ij}$ in each of these and (1) and (2) imply that $a_{il} \neq -a_{ij} \neq a_{im}$. 4 and 7 do not occur since here no a_{ij} can be zero.

Thus from the three fundamental theorems we can write down the transformations which leave the s of any semi-reduced form unaltered and these are one of four types multiplied on the left by certain automorphs if every a_{ij} is non-positive.

7. **A test for equivalence.** It may easily be established that if φ is semi-reduced, permissible transformations A, B, C and D permute the non-vanishing coefficients of φ when no coefficient is positive and the transformations listed in paragraph 6 change, except in order, not more than six[9] non-vanishing coefficients of φ. To test the equivalence of the forms f and f' use transformations A and B to make them semi-reduced. Call φ and φ' the φ functions of their semi-reduced equivalents. φ and φ' *can* be equivalent only if the following are *all* true:

1. The s of φ is equal to the s of φ'.

2. The number of positive coefficients of φ is the same as the number of positive coefficients of φ'.

3a. If no coefficient of φ is positive, the coefficients of φ are the coefficients of φ' in some order.

3b. All but six of the coefficients of φ are, in some order, all but six of the coefficients of φ'.

If 1, 2 and 3 all hold, one can use fundamental theorems 1, 2 and 3 to establish or disestablish equivalence.

CORNELL UNIVERSITY.

---

[8] By a proof which is not long f may be shown to be, under these conditions, positive if its determinant is not zero.

[9] The transformations change not more than four except when a_{ij}, $a_{lm} > 0$ and condition 3 holds.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

PROOF BY CASES IN FORMAL LOGIC

BY S. C. KLEENE

(Received May 8, 1933, Revised March 28, 1934)

1. **Introduction.** By a *logic* or *system of formal logic* we shall mean the totality of formulas consisting of the formulas of a given set (the *formal axioms*) and all the formulas derivable from them by the use of certain formal rules (the *rules of procedure*). The formulas of the logic shall be said to be *true* or *provable*.

We are concerned primarily with the case in which the formal axioms and the rules of procedure are those proposed by A. Church in these annals.[1] In the development of this system the question arises of whether there is any process of formal proof analogous to the intuitive method of proof by cases. (The answer is not immediate, since there is no evident means for proving the implication $\mathbf{F}(\mathbf{x}) \supset_{\mathbf{x}} \mathbf{G}(\mathbf{x})$ between two functions $\mathbf{F}$ and $\mathbf{G}$, except when $\Sigma\mathbf{x}\cdot\mathbf{F}(\mathbf{x})$ can be proved. Consequently direct application of Axiom 25 is not in general available for effecting a proof by cases.) An affirmative answer can be given from the considerations of this paper.

Only certain of the axioms listed by Church are required for our results. When studying the logic based on those axioms only, the provability of formulas in systems obtained by adding other formal axioms is relevant, in view of Theorem I and Corollaries,[2] and like theorems below. Church's terminology and rules of procedure do not lend themselves to a statement and proof of theorems about the provability of formulas in such a form that they apply to all such systems. Hence we introduce slight changes.

Formal symbol was defined by Church.[3] The term *expression* or *formula* we make precise by defining it to denote a finite sequence of formal symbols. One expression shall be said to *occur in*, be *contained in*, or be a *part of* another, if the symbols of the one sequence occur in the same order and contiguously in the other. If the expression $\mathbf{N}$ occurs one or more times in $\mathbf{M}$, we may use the term *occurrence* to denote a particular part of $\mathbf{M}$ of the form $\mathbf{N}$. When $\mathbf{T}$ denotes a particular part of $\mathbf{M}$, we may say that $\mathbf{T}$ is *in* $\mathbf{M}$.

Any formal symbol other than $\{, \}, (,), \lambda, [,]$ shall be a *proper symbol*.

---

[1] A. Church, A set of postulates for the foundation of logic, Vol. 33 (1932), pp. 346–366, and Vol. 34 (1933) pp. 839–864. Arabic numerals will be used to refer to the axioms listed on p. 841, 1933.

[2] Church, 1932, p. 358 and p. 366. These are examples of metamathematical theorems or theorems about the logic, as distinguished from formal theorems or proved formulas of the logic.

[3] 1932, p. 351.

A formula will be called *well-formed*, and an occurrence of a proper symbol in a formula will be said to be an occurrence as a *free* symbol, in accordance with the following rules, and those only: (1) a formula consisting of a single proper symbol **x** is well-formed, and the occurrence of **x** in the formula **x** is an occurrence as a free symbol, (2) if **M** and **N** are well-formed, {**M**}(**N**) is well-formed, and the occurrences of proper symbols as free symbols in **M** and in **N** are occurrences as free symbols in {**M**}(**N**), (3) if **R** is well-formed, and **x** occurs as a free symbol in **R**, λ**x**[**R**] is well-formed, and the occurrences of proper symbols other than **x** as free symbols in **R** are occurrences as free symbols in λ**x**[**R**]. An occurrence of a proper symbol in a formula which is not an occurrence as a free symbol will be called an occurrence as a *bound* symbol. By free a (bound) symbol of **K**, we shall mean a proper symbol which occurs as a free (bound) symbol in **K**.

The axioms of Church's list are well-formed.

Our rules of procedure shall be I-V[4] with "well-formed," "free" and "bound" read in the present sense, "variable" replaced by "proper symbol," and "occurrence of $S^{\mathbf{x}}_{\mathbf{N}}\mathbf{M}$ | in **J**" in III replaced by "occurrence of $S^{\mathbf{x}}_{\mathbf{N}}\mathbf{M}$ | in **J** not immediately following an occurrence of λ."

A proper symbol shall be a *constant symbol* if it is a free symbol of an axiom; otherwise it shall be a *variable*.

Each theorem shall be valid for any logic whose rules of procedure are I-V *and whose formal axioms are well-formed and include* 1, 3–11, 14–16.

This program will be satisfied for metamathematical theorems if their proofs are phrased with reference to any logic of the sort indicated. (Then if variables are employed, letters must be used to represent them ambiguously, since the axiom list is not explicit.) The program will be satisfied for formal theorems if they are shown to be provable when the axioms are 1, 3–11, 14–16.

It is easily shown that, when the axioms are well-formed, the effect of the changes in the rules of procedure is the addition of Π, Σ, &, ∼, ι and A to the set of symbols which may be used as bound symbols in the statement of provable propositions.[5] Theorem I and its corollaries remain true under the new terminology and generality of interpretation.

2. Well-formed expressions.[6] We take note of certain laws concerning the relations among the signs of inclusion in expressions generated by repeated operations of association; and apply them to the well-formed expressions.

---

[4] Church, 1932, p. 355–6. The Roman numerals I–V will henceforth be used to refer to the modified forms of these rules which we are now adopting.

[5] There is also the change resulting from the fact that an occurrence of a proper symbol **x** in a part λ**x**[**M**] of a well-formed expression **K** is now not necessarily an occurrence in **K** as a bound symbol, unless the brackets about **M** "belong together" in a certain sense, which is secured by the condition that **M** be well-formed; and the specified change in Rule III. These two changes merely correct misstatements of what was previously intended.

[6] Some theorems concerning well-formed expressions, similar to theorems of this section, were presented by Church in a course of lectures at Princeton, Nov. 1931.

Given a linear series of elements (*points*) of any sort, we designate by the term *interval* a subset of the elements consisting of all elements between and including a given pair (the *left* and *right endpoints*).

2I. *Given two sets of m points, a_i and b_j, all distinct, in a linear series, if it is possible to pair the a_i one-to-one with the b_j as left and right endpoints, respectively, of a set of intervals such that the pairs of endpoints of each two intervals do not separate each other, then it is possible in only one way.*

The theorem is trivial if $m = 1$. Assuming it true for sets of p points, it follows for sets of $p + 1$ points thus: In the given set of $p + 1$ intervals, there must exist an interval (a_k, b_l) which contains no other. This interval must belong to any other like set which can be constructed from the two sets of $p + 1$ points, for if a_k were paired with some other point b_r $(r \neq l)$, then there would exist no a_s, with which b_l could be paired, such that (a_s, b_l) and (a_k, b_r) do not separate each other. Since (a_k, b_l) thus belongs to both sets of $p + 1$ intervals, the sets of p points $a_t (t \neq k)$ and $b_u (u \neq l)$ come under the hypothesis of the induction, and hence the two sets of $p + 1$ intervals are identical.

2II. *If an interval in the set of a_i and b_j, as well as the whole set, admits a pairing of the type considered in* 2I, *then the pairing in the interval is a part of the pairing in the whole set.*

This may be established in the same manner as 2I.

2III. *In a well-formed expression* **K**, *there exists a unique pairing* (1) *of the occurrences of the symbols* { *with the occurrences of the symbols* }, (2) (*with*), (3) { *with*), (4) λ *with*[, (5) [*with*], (6) λ *with*], (7) { *and* λ *collectively with*) *and*] *collectively, etc., as left and right signs of inclusion, respectively, of parts of* **K** *such that each two parts are either non-overlapping or nested one in the other. Moreover, if such a pairing exists in a part of* **K**, *it is a part of the pairing in* **K**.

List the occurrences of (1) { and }, (2) (and), $\cdots$, as a_i and b_j, respectively. The "known" construction of **K** from proper symbols, which exists according to the definition of *well-formed*, leads us to one pairing of a_i and b_j in which no two pairs separate each other. Then 2I and 2II apply. An a_i and b_j thus paired will be termed *corresponding*.

2IV. *If* **K** *is well-formed, a necessary and sufficient condition that a part* **N** *of* **K** *which occurs between* { *and* }, *or* (*and*), *or* λ *and* [, *or* [*and*], *or begins with* { *and terminates with*), *or begins with* λ *and terminates with*], *be well-formed is that the* { *and* }, *or* (*and*), *or* λ *and* [, *or* [*and*], *or* { *and*), *or* λ *and*], *as the case may be, correspond.*

The case of braces is typical. If **N** is well-formed, there is a correspondence of braces in **N**. Hence there is a correspondence of braces in the part $\{_a\mathbf{N}\}_b$ of **K** in which a and b correspond. The sufficiency of the condition follows from the known construction of **K** as a well-formed expression, since each pair of corresponding braces when introduced encloses a well-formed expression.

2V. *A well-formed part* **N** *of a well-formed expression* **M** *either is* **M**, *or is a proper symbol enclosed between* λ *and* [, *or occurs enclosed by braces, parentheses or brackets.*

Each well-formed part of **M** which is a single symbol enters in the known construction of **M** in one of the situations listed by the theorem. If **N** consists of more than one symbol, then by the definition of a well-formed expression, **N** either begins with { and ends with), or begins with λ and ends with]. Then, by 2IV, this { and), or λ and], correspond; and in the known construction of the correspondence of {'s and)'s, and of λ's and]'s, in **M**, each corresponding pair interior to **M** is enclosed by braces, parentheses, or brackets.

2VI. *If* $\{\mathbf{M}\}(\mathbf{N})$ *and either* **M** *or* **N** *is well-formed, then both* **M** *and* **N** *are well-formed.*

2VII. *A well-formed expression can be of the form* $\{\mathbf{M}\}(\mathbf{N})$, *where* **M** (*or* **N**) *is well-formed, in only one way.*

Proofs. If **M** is well-formed, we suppose $\{_a\mathbf{M}\}_b(\mathbf{N})$ also written $\{_a\mathbf{M}'\}_{b'}(\mathbf{N}')$ where $\mathbf{M}'$ is well-formed. Then there is a correspondence of braces in $\{_a\mathbf{M}\}_b$ in which a and b correspond, and in $\{_a\mathbf{M}'\}_{b'}$ in which a, b' correspond. Hence b and b' are the same occurrence of }; and **M** and $\mathbf{M}'$, also **N** and $\mathbf{N}'$, are identical. Similarly, if **N** is well-formed. But from the definition, we know that a well-formed expression whose first symbol is { is of the form $\{\mathbf{M}\}(\mathbf{N})$ where both **M** and **N** are well-formed.

2VIII. *If* **K**, **P**, *and* **Q** *are well-formed, and* **P** *and* **Q** *are parts of* **K**, *then either* **P** *is in* **Q**, *or* **Q** *is in* **P**, *or* **P** *and* **Q** *have no part in common.*

2IX. *Two distinct occurrences of a well-formed expression* **P** *in a well-formed expression* **K** *have no part in common.*

2X. *If* **M**, **N**, *and* **P** *are well-formed, and* **P** *is a part of* $\{\mathbf{M}\}$ $(\mathbf{N})$, *then either* **P** *is* $\{\mathbf{M}\}(\mathbf{N})$, *or* **P** *is in* **M**, *or* **P** *is in* **N**.

Proofs. If **P** and **Q** had a part in common without either containing the other, then they would each have more than one symbol, and hence begin with a { or λ and end with a corresponding) or], contradicting the property of the correspondence 2III (7) that pairs of corresponding symbols do not separate each other. Thus 2VIII is established. 2IX and 2X are corollaries.

2XI. *If* $\lambda\mathbf{x}\,[\mathbf{R}]$ *and either* **x** *or* **R** *are well-formed, then both* **x** *and* **R** *are well-formed, and* **x** *occurs in* **M** *as a free symbol.*

2XII. *If* $\lambda\mathbf{x}\,[\mathbf{R}]$, **R** (*or* **x**), *and* **P** *are well-formed, and* **P** *occurs in* $\lambda\mathbf{x}\,[\mathbf{R}]$, *then either* **P** *is* $\lambda\mathbf{x}\,[\mathbf{R}]$, *or* **P** *is* **x**, *or* **P** *occurs in* **R**.

2XIII. *An occurrence of a proper symbol* **x** *in a well-formed expression* **K** *is an occurrence as a bound symbol, or as a free symbol, according as it is, or is not, an occurrence in a well-formed part of* **K** *of the form* $\lambda\mathbf{x}\,[\mathbf{R}]$.

2XIV. *If* **L** *is well-formed, if the proper symbol* **x** *does not occur in* **L** *as a free symbol, and if the proper symbol* **y** *does not occur in* **L**, *then* $S^{\mathbf{x}}_{\mathbf{y}}\mathbf{L}\mid$ *is well-formed, and* **L** *and* $S^{\mathbf{x}}_{\mathbf{y}}\mathbf{L}\mid$ *have the same free symbols.*

2XV. *If* **M** *and* **N** *are well-formed, if the proper symbol* **x** *occurs in* **M**, *and if the bound symbols of* **M** *are distinct both from* **x** *and from the free symbols of* **N**, *then* $S^{\mathbf{x}}_{\mathbf{N}}\mathbf{M}\mid$ *and* $\{\lambda\mathbf{x}\cdot\mathbf{M}\}(\mathbf{N})$ *are well-formed and have the same free symbols.*

The proofs of 2XI and 2XII are similar to those of 2VI and 2X, respectively. To establish 2XIII, we infer from the definition of *bound* that if an occurrence

of $\mathbf{x}$ in $\mathbf{K}$ is an occurrence as a bound symbol, then in every construction of $\mathbf{K}$ from proper symbols as a well-formed expression, there must enter a well-formed part $\lambda\mathbf{x}\,[\mathbf{R}]$ containing the occurrence of $\mathbf{x}$. Moreover, were there any part of $\mathbf{K}$ of the latter form which did not enter in the course of every such construction, the correspondence of λ's and]'s in that part would not be a part of the correspondence in $\mathbf{K}$, contradicting 2III(6). 2XIV and 2XV are corollaries.

2XVI. *If* $\mathbf{J}$, $\mathbf{X}$, *and* $\mathbf{X}'$ *are well-formed, and if* $\mathbf{X}$ *and* $\mathbf{X}'$ *have the same free symbols, then* $\mathbf{J}'$, *the expression obtained by substituting* $\mathbf{X}'$ *for a particular occurrence of* $\mathbf{X}$ *in* $\mathbf{J}$ *not immediately following an occurrence of* λ, *is well-formed.*

The occurrence of $\mathbf{X}$ in question in 2XVI enters in the course of the known construction of $\mathbf{J}$ (otherwise by utilizing 2V and 2IV we could contradict the last clause of 2III). $\mathbf{X}'$ can be constructed and introduced instead of $\mathbf{X}$, and a construction of $\mathbf{J}'$ then completed by steps similar to those of the remainder of the construction of $\mathbf{J}$.

2XVII. *Every provable formula is well-formed.*

We are admitting only well-formed expressions as axioms. An application of IV or V to well-formed expressions yields a well-formed expression; and the same is true of I, by 2XIV, 2XIII and 2V (by which, if $\mathbf{x}$ occurs in $\mathbf{L}$, $\mathbf{L}$ does not occur in $\mathbf{J}$ immediately following λ), and 2XVI; and of II and III, by 2XV, 2V and 2XVI.

A part $\mathbf{P}$ of $\mathbf{K}$ shall be termed *free* if each occurrence of a proper symbol in $\mathbf{P}$ as a free symbol is an occurrence in $\mathbf{K}$ as a free symbol; otherwise $\mathbf{P}$ shall be *bound*. If $\mathbf{P}$ is a proper symbol, this definition reduces to the one of §1.

We may without ambiguity generalize the substitution notation of Church[7] by using $S^{\mathbf{X}}_{\mathbf{Y}}\mathbf{U}\,|$ to represent the expression which is obtained by replacing each occurrence (if any) of $\mathbf{X}$ in $\mathbf{U}$ by $\mathbf{Y}$, when $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{U}$ are any well-formed expressions. (Cf. 2IX.) $S^{\mathbf{X}_1}_{\mathbf{Y}_1}\cdots S^{\mathbf{X}_n}_{\mathbf{Y}_n}\mathbf{U}\,|\cdots|$ may be abbreviated $S^{\mathbf{X}_1}_{\mathbf{Y}_1}\!:\!:\!:\,^{\mathbf{X}_n}_{\mathbf{Y}_n}\mathbf{U}\,|$.

2XVIII. *An occurrence of a well-formed expression* $\mathbf{X}$ *in a well-formed expression* $\mathbf{K}$ *is an occurrence as a bound expression, or as a free expression, according as it is, or is not, an occurrence in a well-formed part of* $\mathbf{K}$ *of the form* $\lambda\mathbf{x}\,[\mathbf{R}]$ *where* $\mathbf{x}$ *occurs in* $\mathbf{X}$ *as a free symbol.*

2XIX. *If* $\mathbf{K}$, $\mathbf{X}$, *and* $\mathbf{X}'$ *are well-formed, then* $\mathbf{K}'$, *the expression obtained from* $\mathbf{K}$ *by substituting* $\mathbf{X}'$ *for a particular occurrence of* $\mathbf{X}$ *in* $\mathbf{K}$ *as a free expression, is well-formed.*

2XX. *If* $\mathbf{M}$, $\mathbf{X}$, *and* $\mathbf{N}$ *are well-formed, and* $\mathbf{X}$ *does not occur in* $\mathbf{M}$ *as a bound expression, then* $S^{\mathbf{X}}_{\mathbf{N}}\mathbf{M}\,|$ *is well-formed.*

Theorem 2XVIII is a corollary of 2XIII. 2XIX is obtained in the same manner as 2XVI; and 2XX is a corollary of 2XIX.

3. **Abbreviation of formulas.** We reserve the **heavy-type** letters to represent undetermined formulas in metamathematical discussions; and adopt the conventions that, henceforth, (1) *each heavy-type letter shall represent a well-*

[7] 1932, p. 350.

formed expression, and (2) *each set of symbols standing apart which contains a heavy-type letter shall represent a well-formed expression.* Our operations with heavy-type letters under these conventions involve tacit use of the theorems of §2.

For example, if $\lambda\mathbf{x}\,[\mathbf{R}]$ is mentioned, it is supposed, in accordance with (1) and (2), that $\mathbf{x}$, $\mathbf{R}$, and $\lambda\mathbf{x}\,[\mathbf{R}]$ represent well-formed expressions. It follows by 2XI that $\mathbf{x}$ represents a proper symbol occurring in the expression represented by $\mathbf{R}$ as a free symbol. Similarly, when mentioning $\mathbf{M} \supset_{\mathbf{x}} \mathbf{N}$ (i.e. $\{\{\Pi\}\ (\lambda\mathbf{x}\,[\mathbf{M}])\}\ (\lambda\mathbf{x}\,[\mathbf{N}])$), we suppose that $\mathbf{M}$, $\mathbf{x}$, $\mathbf{N}$, and $\mathbf{M} \supset_{\mathbf{x}} \mathbf{N}$ "are" well-formed. It follows, by means of 2III and 2IV, that $\lambda\mathbf{x}\,[\mathbf{M}]$ and $\lambda\mathbf{x}\,[\mathbf{N}]$ are well-formed; and hence, by 2XI, that $\mathbf{x}$ is a free symbol of both $\mathbf{M}$ and $\mathbf{N}$.

We also employ specific abbreviations, including those of Church, 1932, §6. The discussion of formulas represented by abbreviations is facilitated by imposing the restriction on our methods of specific abbreviation that (3) *each occurrence of a free symbol other than* Π, Σ, $\&$, $\sim$, ι, A *in the formula shall be represented by an occurrence of that symbol in the abbreviation, unless the abbreviation contains a heavy-type letter.* Our rules of abbreviation make it easy to determine from an abbreviation whether a given part which represents a proper symbol represents a set of free occurrences, or a set of bound occurrences, of the symbol.

In applying the conventions concerning omission of braces and brackets, we shall sometimes treat heavy-type letters as though they were single proper symbols.

The symbol "$\rightarrow$" indicating definition may be read "stands for," "stand for," "to stand for," etc., according to the context. We use the definition of $=$ given by Church (1933, p. 844). German letters are used for more or less temporary definitions.

The abbreviation of $\lambda\mathbf{x}_1\,[\,\cdots\,\lambda\mathbf{x}_n\,[\mathbf{R}]\,\cdots\,]$ to $\lambda\mathbf{x}_1 \cdots \lambda\mathbf{x}_n \cdot \mathbf{R}$ will be carried further to $\lambda\mathbf{x}_1 \cdots \mathbf{x}_n \cdot \mathbf{R}$. Similarly, $\Sigma\mathbf{x}_1 \cdots \Sigma\mathbf{x}_n \cdot \mathbf{M}$ will be shortened to $\Sigma\mathbf{x}_1 \cdots \mathbf{x}_n \cdot \mathbf{M}$; etc.

$[[\mathbf{P}]\,[\mathbf{Q}]]\,[\mathbf{R}]$ will be abbreviated further to $[\mathbf{P}]\,[\mathbf{Q}]\,[\mathbf{R}]$. Similarly in the case of products and sums of other sorts.

4. **The symbol $\vdash$.**[8] "$\mathbf{P}, \mathbf{Q}, \mathbf{R}, \cdots \vdash \mathbf{T}$" shall mean that $\mathbf{T}$ is derivable from $\mathbf{P}, \mathbf{Q}, \mathbf{R}, \cdots$ by zero or more applications of our rules of procedure (I–V). To this definition we add a convention which makes the meaning of $\vdash$ depend also on the set of formal axioms: such of $\mathbf{P}, \mathbf{Q}, \mathbf{R}, \cdots$ in "$\mathbf{P}, \mathbf{Q}, \mathbf{R}, \cdots \vdash \mathbf{T}$" as are provable *may* be omitted.

Thus $\mathbf{M}, \mathbf{N}, \cdots \vdash \mathbf{T}$ is the assertion that $\mathbf{T}$ is provable as a consequence of $\mathbf{M}, \mathbf{N}, \cdots$ in the sense that $\mathbf{T}$ could be made provable by adding $\mathbf{M}, \mathbf{N}, \cdots$ to the list of axioms;[9] and $\vdash \mathbf{T}$ is the assertion that $\mathbf{T}$ is provable. (The symbol $\vdash$ in the phrase "$\vdash \mathbf{T}$" will often be omitted.[10] However "$\mathbf{T}$" standing by itself

[8] Our definition of this symbol was suggested by that of J. B. Rosser.

[9] Cf. Church, 1932, p. 357.

[10] Cf. Church, 1932, p. 350.

sometimes means merely that $\mathbf{M}, \mathbf{N}, \cdots \vdash \mathbf{T}$, where $\mathbf{M}, \mathbf{N}, \cdots$ have been previously "assumed." "$\vdash$" may be read "yields.")

Various obvious relations involving $\vdash$ will be used tacitly.

In view of Axioms 14–16, if $\mathbf{A}_1, \cdots, \mathbf{A}_n \vdash \mathbf{C}$, then $\mathbf{A}_1 \cdots \mathbf{A}_n \vdash \mathbf{C}$, and conversely; and if $\mathbf{A}_1, \cdots, \mathbf{A}_n \vdash \mathbf{B}_i$ $(i = 1, \cdots, m)$, then $\mathbf{A}_1, \cdots, \mathbf{A}_n \vdash \mathbf{B}_1 \cdots \mathbf{B}_m$.

We have arranged that the theorems in which $\vdash$ appears should contain no other terms whose definition is relative to the logic except *variable*. Accordingly, in view of the program announced in §1, we have the following rule:

Each theorem in which $\vdash$ *appears remains valid when* "$\vdash$" *is replaced by* "$\mathbf{X}, \mathbf{Y}, \cdots \vdash$," *and "variable" by "variable distinct from the free variables of* $\mathbf{X}, \mathbf{Y}, \cdots$."

We may apply this rule to Theorem I and its corollaries, supposing them stated in the $\vdash$ notation.

5. **Conversion, notes.** We employ the terms *conversion* and *convertible* with the understanding that I–III represent the new Rules I–III.[11] "$\mathbf{M}$ is convertible into $\mathbf{N}$" will be abbreviated to "$\mathbf{M}$ conv $\mathbf{N}$."

We use tacitly the reflexive, symmetric, and transitive properties of the relation $\mathbf{M}$ *conv* $\mathbf{N}$, and the property that if $\mathbf{M}$ conv $\mathbf{N}$ and $\mathbf{J}'$ is the result of substituting $\mathbf{N}$ for an occurrence of $\mathbf{M}$ in $\mathbf{J}$ not immediately following λ, $\mathbf{J}$ conv $\mathbf{J}'$.

5I. *Suppose that* $\mathbf{x}$ *occurs in* $\mathbf{P}$ *as a free symbol, and let* $\mathbf{P}'$ *denote the expression obtained by substituting* $\mathbf{Q}$ *for the free occurrences of* $\mathbf{x}$ *in* $\mathbf{P}$. *If the resulting occurrences of* $\mathbf{Q}$ *in* $\mathbf{P}'$ *are free, then* $\{\lambda\mathbf{x}\cdot\mathbf{P}\}(\mathbf{Q})$ *conv* $\mathbf{P}'$.

For, under these conditions, it is possible by applications of Rule I to convert $\mathbf{P}$ and $\mathbf{P}'$ into expressions $\mathbf{R}$ and $\mathbf{R}'$, respectively, such that $\{\lambda\mathbf{x}\cdot\mathbf{R}\}(\mathbf{Q})$ yields $\mathbf{R}'$ by an application of Rule II.

$\mathbf{M}'$ shall be a *normal form* of $\mathbf{M}$, if $\mathbf{M}$ conv $\mathbf{M}'$ and $\mathbf{M}'$ contains no part of the form $\{\lambda\mathbf{x}\cdot\mathbf{R}\}(\mathbf{S})$. The assertion "$\mathbf{M}$ conv $\mathbf{N}$" can often be verified by converting $\mathbf{M}$ into a normal form $\mathbf{M}'$, and $\mathbf{N}$ into a normal form $\mathbf{N}'$, by applications of I and II, and finding that $\mathbf{M}'$ and $\mathbf{N}'$, if not identical, can be made so by applications of I.

If $\mathbf{M}$ conv $\mathbf{M}'$, then $\mathbf{M} \vdash \mathbf{M}'$ and $\mathbf{M}' \vdash \mathbf{M}$. Hence it is sufficient for the proof that $\mathbf{M}_1, \cdots, \mathbf{M}_n \vdash \mathbf{N}$ to show that $\mathbf{M}'_1, \cdots, \mathbf{M}'_n \vdash \mathbf{N}'$, where $\mathbf{M}'_1$ conv $\mathbf{M}_1, \cdots, \mathbf{M}'_n$ conv $\mathbf{M}_n$, and $\mathbf{N}'$ conv $\mathbf{N}$. We often use this principle tacitly. For example, we may pass from $\mathbf{m} \supset_{\mathbf{x}} \mathbf{n}$, and the result $\mathbf{M}$ of substituting $\mathbf{X}$ for the free occurrences of $\mathbf{x}$ in $\mathbf{m}$, to the result $\mathbf{N}$ of substituting $\mathbf{X}$ for the free occurrences of $\mathbf{x}$ in $\mathbf{n}$, *when* the resulting occurrences of $\mathbf{X}$ in $\mathbf{M}$ and in $\mathbf{N}$ are free (cf. 5I and Rule V).

5II. *If* $\mathbf{X}$ *occurs in* $\mathbf{M}$ *as a free expression, then* $\mathbf{M} \vdash E(\mathbf{X})$ *(and* $E(\mathbf{M}) \vdash E(\mathbf{X})$*).*

For, by successive applications of 5I, $\mathbf{M}$ conv $\{\lambda\mathbf{x}\cdot\mathbf{M}'\}(\mathbf{X})$, conv $\{\lambda\mathbf{f}\cdot\mathbf{f}(\mathbf{X})\}(\lambda\mathbf{x}\cdot\mathbf{M}')$, for suitably chosen $\mathbf{x}$, $\mathbf{M}'$ and $\mathbf{f}$. From the last expression, by IV, $\Sigma\mathbf{f}\cdot\mathbf{f}(\mathbf{X})$; and thence, by conversion, $E(\mathbf{X})$.

5III. *If* $\mathbf{x}_1, \cdots, \mathbf{x}_n$ *are variables,* $\{\lambda\mathbf{x}_1 \cdots \mathbf{x}_n\cdot\mathbf{M}\}(\mathbf{X}_1, \cdots, \mathbf{X}_n) \vdash \Sigma\mathbf{x}_1 \cdots \mathbf{x}_n\cdot\mathbf{M}$.

[11] Church, 1932, p. 357.

5IV. *If the set of the variables* $\mathbf{x}_1, \cdots, \mathbf{x}_n$ *which occur in* $\mathbf{N}$ *as free symbols is* $\mathbf{x}_{\sigma_1}, \mathbf{x}_{\sigma_2}, \cdots, \mathbf{x}_n$, *then* $\{\lambda\mathbf{x}_1 \cdots \mathbf{x}_n \cdot \mathbf{M}\}(\mathbf{X}_1, \cdots, \mathbf{X}_n), \mathbf{M} \supset_{\mathbf{x}_1 \cdots \mathbf{x}_n} \cdot \mathbf{N} \vdash \{\lambda\mathbf{x}_{\sigma_1}\mathbf{x}_{\sigma_2} \cdots \mathbf{x}_n \cdot \mathbf{N}\}(\mathbf{X}_{\sigma_1}, \mathbf{X}_{\sigma_2}, \cdots, \mathbf{X}_n)$.

For it can be shown by induction with respect to n that $\Sigma\mathbf{x}_1 \cdots \mathbf{x}_n \cdot \mathbf{M}$ can be derived from $\{\lambda\mathbf{x}_1 \cdots \mathbf{x}_n \cdot \mathbf{M}\}(\mathbf{X}_1, \cdots, \mathbf{X}_n)$ by conversions (using I and 5I) and applications of IV; and that $\{\lambda\mathbf{x}_{\sigma_1} \cdots \mathbf{x}_n \cdot \mathbf{N}\}(\mathbf{X}_{\sigma_1}, \cdots, \mathbf{X}_n)$ can be derived from $\{\lambda\mathbf{x}_1 \cdots \mathbf{x}_n \cdot \mathbf{M}\}(\mathbf{X}_1, \cdots, \mathbf{X}_n)$ and $\mathbf{M} \supset_{\mathbf{x}_1 \cdots \mathbf{x}_n} \cdot \mathbf{N}$ by conversions and applications of IV and V.

5V. *If* $\mathbf{P}, \mathbf{Q}, \cdots \vdash \mathbf{T}$, *and* $\mathbf{x}$ *occurs in* $\mathbf{T}$ *as a free variable, then* $\mathbf{x}$ *occurs in at least one of* $\mathbf{P}, \mathbf{Q}, \cdots$ *as a free variable.*

COROLLARY. *No provable formula contains free variables.*

5VI. *If* $\mathbf{B}$ *conv* $\mathbf{C}$, *then* $\mathbf{B}$ *and* $\mathbf{C}$ *have the same free symbols.*

For an application of one of the rules I–V to an expression $\mathbf{A}$ (expressions $\mathbf{A}$ and $\mathbf{B}$) yields an expression $\mathbf{C}$ whose free symbols (other than Σ in the case of IV) are included among those of $\mathbf{A}$ ($\mathbf{A}$ and $\mathbf{B}$). Hence the free symbols of $\mathbf{T}$ are included among those of $\mathbf{P}, \mathbf{Q}, \cdots$ and the axioms (Σ is a free symbol of the axioms); and the free symbols of $\mathbf{C}$ are included among those of $\mathbf{B}$, and *vice versa.*

6. **Combinations.** We adopt the definitions

$$I \to \lambda x \cdot x, \qquad J \to \lambda wxyz \cdot w(x, w(z, y));$$

and define *combination* by complete induction thus: (1) I, J, and proper symbols are combinations. (2) If $\mathbf{M}$ and $\mathbf{N}$ are combinations, then $\{\mathbf{M}\}(\mathbf{N})$ is a combination.

6I. *A combination is well-formed.*

6II. *If* $\mathbf{M}(\mathbf{N})$ *is a combination, then* $\mathbf{M}$ *and* $\mathbf{N}$ *are combinations.* (Cf. 2VII.)

By a *term* of a combination $\mathbf{C}$, we shall mean a free occurrence of I, J, or a proper symbol, in $\mathbf{C}$.

6III. *A combination* $\mathbf{C}$ *of* k *terms consists of a unique linear series of its* k *terms, interspersed with braces and parentheses.*

For by definition, $\mathbf{C}$ can be constructed from I, J, and proper symbols in a manner which shows that $\mathbf{C}$ consists of a certain linear series of terms of $\mathbf{C}$, interspersed with braces and parentheses. It follows from 2VIII that either every term of $\mathbf{C}$ occurs in this series, or some term of $\mathbf{C}$ is a part, not the whole, of another term. The latter is impossible, since any term composed of more than one symbol is either an I or a J, and hence has no free part except itself.

6IV. *If the last term of a combination be replaced by* $\mathbf{A}$, *and the others (if any) by* I, *the resulting expression is convertible into* $\mathbf{A}$.

This theorem is provable by induction with respect to the number of terms of the given combination, using the following facts: In view of the definitions, a combination $\mathbf{C}$ of $n + 1$ terms is of the form $\mathbf{C}_1(\mathbf{C}_2)$ where $\mathbf{C}_1$ and $\mathbf{C}_2$ are combinations. Then, by 6III, the series of terms of $\mathbf{C}$ is the series of terms of $\mathbf{C}_1$ followed by that of $\mathbf{C}_2$. $I(\mathbf{A})$ conv $\mathbf{A}$.

The theory of combinatory logic, developed by M. Schönfinkel and H. B. Curry, has been brought into relation to the theory of the present system by J. B. Rosser.[12] The following two theorems result from Rosser's work:

6V. *Given* $\mathbf{M}$, *a combination* $\mathbf{M}'$ *can be found such that* $\mathbf{M}'$ *conv* $\mathbf{M}$.

6VI. *If* $\mathbf{K}_1, \cdots, \mathbf{K}_n$ *and* $\mathbf{L}$ *are combinations, and* $\mathbf{L}$ *is derivable from* $\mathbf{K}_1, \cdots, \mathbf{K}_n$ *by Rules* I–V, *then* $\mathbf{L}$ *is derivable from* $\mathbf{K}_1, \cdots, \mathbf{K}_n$ *and certain provable combinations* $\mathfrak{A}_1, \cdots, \mathfrak{A}_{16}$ *by means of the following rules:*

R1. *If* $\Pi(\mathbf{f}, \mathbf{g})$ *and* $\mathbf{f}(\mathbf{p})$, *then* $\mathbf{g}(\mathbf{p})$.

R2. *If* $\mathbf{f}(\mathbf{p})$, *then* $\Sigma(\mathbf{f})$.

R3. *If* $I(\mathbf{p})$, *then* $\mathbf{p}$.

R4. *If* $\mathbf{p}$, *then* $I(\mathbf{p})$.

R5. *If* $\mathbf{f}(I(\mathbf{p}, \mathbf{q}))$, *then* $\mathbf{f}(\mathbf{p}(\mathbf{q}))$.

R6. *If* $\mathbf{f}(\mathbf{p}(\mathbf{q}))$, *then* $\mathbf{f}(I(\mathbf{p}, \mathbf{q}))$.

R7. *If* $\mathbf{f}(J(\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}))$, *then* $\mathbf{f}(\mathbf{p}(\mathbf{q}, \mathbf{p}(\mathbf{s}, \mathbf{r})))$.

R8. *If* $\mathbf{f}(\mathbf{p}(\mathbf{q}, \mathbf{p}(\mathbf{s}, \mathbf{r})))$, *then* $\mathbf{f}(J(\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}))$.

All the formulas occurring in this derivation are combinations.

The last clause is a consequence of 6II.

6VII. *If* $\mathbf{A}_1, \cdots, \mathbf{A}_n$ *contain no free symbols, then there can be found expressions* $\mathbf{C}_1, \cdots, \mathbf{C}_n, \mathbf{H}$ *such that* $\mathbf{C}_i(I)$ *conv* $\mathbf{A}_i$ *and* $\mathbf{H}(\mathbf{C}_i)$ *conv* I $(i = 1, \cdots, n)$.

By 6V, $\mathbf{A}_i$ is convertible into a combination $\mathbf{B}_i$. Let the number of terms of $\mathbf{B}_i$ be m_i, and let m be any integer $\geqq$ the n numbers $m_1, \cdots, m_n$. Let $\mathbf{B}'_i$ represent the combination $I(\cdots I(\mathbf{B}_i)\cdots)$ of m terms, $\mathbf{b}_{i1}, \cdots, \mathbf{b}_{im}$. By 6III, we can without ambiguity define $\mathbf{B}''_i$ to be the expression obtained from $\mathbf{B}'_i$ by substituting $b_1, \cdots, b_m$ for $\mathbf{b}_{i1}, \cdots, \mathbf{b}_{im}$, respectively. Let $\mathbf{C}_i \rightarrow \lambda x \cdot x(\lambda b_1 \cdots b_m \cdot \mathbf{B}''_i, \mathbf{b}_{i1}, \cdots, \mathbf{b}_{im})$. We infer from 5VI, the definition of term, and the hypothesis that $\mathbf{A}_i$ contains no free symbols, that $\mathbf{b}_{ir}$ $(r = 1, \cdots, m)$ is either an I or a J. Now $\mathbf{C}_i(I)$ conv $\{\lambda b_1 \cdots b_m \cdot \mathbf{B}''_i\}$ $(\mathbf{b}_{i1}, \cdots, \mathbf{b}_{im})$, conv $\mathbf{B}'_i$, conv $\mathbf{B}_i$, conv $\mathbf{A}_i$. Let $\mathbf{H} \rightarrow \lambda p \cdot p(\lambda a b_1 \cdots b_m \cdot a(b_1(I, I, I, I), \cdots, b_m(I, I, I, I)))$. $\mathbf{b}_{ir}(I, I, I, I)$ conv I, since both $I(I, I, I, I)$ and $J(I, I, I, I)$ are convertible into I. Hence $\mathbf{H}(\mathbf{C}_i)$ conv $\{\lambda b_1 \cdots b_m \cdot \mathbf{B}''_i\}$ $(\mathbf{b}_{i1}(I, I, I, I), \cdots, \mathbf{b}_{im}(I, I, I, I))$, conv $\{\lambda b_1 \cdots b_m \cdot \mathbf{B}''_i\}$ $(I, \cdots, I)$, conv $S_I^{b_1 \cdots b_m} \mathbf{B}''_i |$, conv I (by 6IV).

---

[12] See M. Schönfinkel, Über die Bausteine der mathematischen Logik, Math. Annalen, Vol. 92 (1924), pp. 305–316; H. B. Curry, An analysis of logical substitution, Am. Journal of Math., Vol. 51 (1929), pp. 363–384; Grundlagen der kombinatorischen Logik, Am. Journal of Math., Vol. 52 (1930), pp. 509–536, 789–834; Some additions to the theory of combinators, Am. Journal of Math., Vol. 54 (1932), pp. 551–558; J. B. Rosser, A mathematical logic without variables, forthcoming in these Annals.

Under Rosser's definitions, I and J are not abbreviations, but undefined symbols of the logic; and there is a difference between the present notation, and the notation used in Rosser's work, for the application of a function to an argument. But there is a one-to-one correspondence between his combinations and the combinations in the present sense, which enables us to apply his theory to our combinations.

7. **The formulas 1 and 2.** We use the definitions

$$1 \rightarrow \lambda f x \cdot f(x), \qquad S \rightarrow \lambda \rho f x \cdot f(\rho(f, x)), \qquad 2 \rightarrow S(1),$$

of Church.[13]

7I. *If* $\mathbf{A}_1$ *and* $\mathbf{A}_2$ *have the same free symbols, then there can be found a formula* $\mathbf{L}$ *such that* $\mathbf{L}(1)$ *conv* $\mathbf{A}_1$ *and* $\mathbf{L}(2)$ *conv* $\mathbf{A}_2$.

Let the free symbols of $\mathbf{A}_1$ be $\mathbf{z}_1, \cdots, \mathbf{z}_m$. Then $\lambda \mathbf{z}_1 \cdots \mathbf{z}_m \cdot \mathbf{A}_i$ $(i = 1, 2)$ has no free symbols; and hence, by 6VII, there can be found formulas $\mathbf{C}_i$, $\mathbf{H}$ such that $\mathbf{C}_i(I)$ conv $\lambda \mathbf{z}_1 \cdots \mathbf{z}_m \cdot \mathbf{A}_i$ and $\mathbf{H}(\mathbf{C}_i)$ conv I. By 5VI, $\mathbf{C}_i$ and $\mathbf{H}$ contain no free symbols. Let $\mathbf{n}$ be a proper symbol distinct from $\mathbf{z}_1, \cdots, \mathbf{z}_m$, and let $\mathbf{L} \rightarrow \lambda \mathbf{n} \cdot \mathbf{n}(\lambda pq \cdot q(p(\lambda b \cdot \mathbf{H}(b, \mathbf{C}_2))), \lambda f \cdot \mathbf{H}(f(\mathbf{C}_1), \mathbf{C}_1), I, I, \mathbf{z}_1, \cdots, \mathbf{z}_m)$. Then $\mathbf{L}(1)$ conv $\{\lambda pq \cdot q(p(\lambda b \cdot \mathbf{H}(b, \mathbf{C}_2)))\}$ $(\lambda f \cdot \mathbf{H}(f(\mathbf{C}_1), \mathbf{C}_1), I, I, \mathbf{z}_1, \cdots, \mathbf{z}_m)$ (using the definition of 1), conv $\{\lambda q \cdot q(\mathbf{H}(\mathbf{H}(\mathbf{C}_1, \mathbf{C}_2), \mathbf{C}_1))\}$ $(I, I, \mathbf{z}_1, \cdots, \mathbf{z}_m)$, conv $\{\lambda q \cdot q(\mathbf{C}_1)\}$ $(I, I, \mathbf{z}_1, \cdots, \mathbf{z}_m)$ (using the conversions of $\mathbf{H}(\mathbf{C}_1)$ and $\mathbf{H}(\mathbf{C}_2)$ into I), conv $\mathbf{C}_1(I, \mathbf{z}_1, \cdots, \mathbf{z}_m)$, conv $\{\lambda \mathbf{z}_1 \cdots \mathbf{z}_m \cdot \mathbf{A}_1\}$ $(\mathbf{z}_1, \cdots, \mathbf{z}_m)$, conv $\mathbf{A}_1$. Similarly, $\mathbf{L}(2)$ conv $\mathbf{L}(\lambda f x \cdot f(f(x)))$, conv $\{\lambda pq \cdot q(p(\lambda b \cdot \mathbf{H}(b, \mathbf{C}_2)))\}$ $(\{\lambda pq \cdot q(p(\lambda b \cdot \mathbf{H}(b, \mathbf{C}_2)))\}$ $(\lambda f \cdot \mathbf{H}(f(\mathbf{C}_1), \mathbf{C}_1)), I, I, \mathbf{z}_1, \cdots, \mathbf{z}_m)$, conv $\{\lambda pq \cdot q(p(\lambda b \cdot \mathbf{H}(b, \mathbf{C}_2)))\}$ $(\lambda q \cdot q(\mathbf{C}_1), I, I, \mathbf{z}_1, \cdots, \mathbf{z}_m)$, conv $\{\lambda q \cdot q(\mathbf{H}(\mathbf{C}_1, \mathbf{C}_2))\}$ $(I, I, \mathbf{z}_1, \cdots, \mathbf{z}_m)$, conv $\{\lambda q \cdot q(\mathbf{C}_2)\}$ $(I, I, \mathbf{z}_1, \cdots, \mathbf{z}_m)$, conv $\mathbf{A}_2$.

Let

$$M \rightarrow \lambda \mu \cdot \phi(1)\phi(2) \supset {}_{\phi}\phi(\mu).$$

Let $\mathbf{A}$ be any formal axiom. By 7I there exists an $\mathbf{F}$ such that $\mathbf{F}(1)$ conv $\mathbf{F}(2)$ conv $\mathbf{A}$. Then $\mathbf{F}(1)\mathbf{F}(2)$, and consequently $\Sigma\phi \cdot \phi(1)\phi(2)$, is provable. Moreover, assuming $\phi(1)\phi(2)$, we can prove both $\phi(1)$ and $\phi(2)$ as consequences. Hence,according to Theorem I, $\phi(1)\phi(2) \supset {}_{\phi}\phi(1)$ and $\phi(1)\phi(2) \supset {}_{\phi}\phi(2)$ are provable. The last two formulas are convertible into $M(1)$ and $M(2)$, respectively. Thus we establish the theorems

7.1. $\qquad \vdash M(1).$

7.2. $\qquad \vdash M(2).$

We adopt the convention that *when an italicized letter is being used to represent the number one (the number two), the letter in heavy-type shall represent the formula* 1 *(the formula* 2*)*.

8. **Representation of a proof by an implication.** In this section we prove

8I. *Let* $\mathbf{x}$ *be any variable distinct from the free variables of* $\mathbf{M}$. *If* $\mathbf{M}(1) \vdash \mathbf{N}$ *and* $\mathbf{M}(2) \vdash \mathbf{N}$, *then there can be found expressions* $\mathbf{M}^*$, 1^*, 2^*, $\mathbf{T}$, $\mathbf{N}^*$, $\mathfrak{R}$ *such that* $\mathbf{M}^*(\mathfrak{R})$ *conv* $\mathbf{M}$, $1^*(\mathfrak{R})$ *conv* 1, $2^*(\mathfrak{R})$ *conv* 2, $\vdash \mathbf{T}(\mathfrak{R})$, $\mathbf{N}^*(\mathfrak{R})$ *conv* $\mathbf{N}$, $E(\mathbf{M}) \vdash \mathbf{M}^*(\mathbf{x}, 1^*(\mathbf{x}))\mathbf{T}(\mathbf{x}) \supset {}_{\mathbf{x}}\mathbf{N}^*(\mathbf{x})$, *and* $E(\mathbf{M}) \vdash \mathbf{M}^*(\mathbf{x}, 2^*(\mathbf{x}))\mathbf{T}(\mathbf{x}) \supset {}_{\mathbf{x}}\mathbf{N}^*(\mathbf{x})$.

Throughout the proof of this theorem, the index j shall take the values 1 and 2, the convention of §7 will be employed, and, furthermore, when j represents 1 (2), $\mathbf{j}^*$ shall represent 1^* (2^*).

[13] 1933, p. 863.

According to the hypothesis, a proof of $\mathbf{N}$ as a consequence of $\mathbf{M}(\mathbf{j})$ can be found. Let $\mathbf{A}_{j1}, \cdots, \mathbf{A}_{j\eta j}$ be a list of the formal axioms used in this proof. According to 6V, combinations $\mathbf{B}$, $\mathbf{j}'$, $\mathbf{B}_{ji}$ $(i = 2, \cdots, \eta_j + 1)$, $\mathbf{N}'$ can be found such that $\mathbf{B}$ conv $\mathbf{M}$, $\mathbf{j}'$ conv $\mathbf{j}$, $\mathbf{B}_{ji}$ conv $\mathbf{A}_{j,i-1}$, $\mathbf{N}'$ conv $\mathbf{N}$. From these conversions and the derivation of $\mathbf{N}$ from $\mathbf{M}(\mathbf{j})$, $\mathbf{A}_{j1}, \cdots, \mathbf{A}_{j\eta j}$ by Rules I–V, we can discover a derivation of $\mathbf{N}'$ from $\mathbf{B}(I(\mathbf{j}'))$, $\mathbf{B}_{j2}, \cdots, \mathbf{B}_{j,\eta j+1}$ by I–V. It follows, by 6VI, that a derivation of $\mathbf{N}'$ from $\mathbf{B}(I(\mathbf{j}'))$, $\mathbf{B}_{j2}, \cdots, \mathbf{B}_{j,\eta j+1}$ and certain provable combinations $\mathfrak{A}_1, \cdots, \mathfrak{A}_{16}$ by Rules R1–R8 can be found. Every formula occurring in this derivation is a combination. Let $\mathbf{B}_j$ be the combination $I(\mathbf{j}')$; $\mathbf{B}_{j1}$ be the combination $\mathbf{B}(\mathbf{B}_j)$; and denote the combinations $\mathfrak{A}_1, \cdots, \mathfrak{A}_{16}$ by $\mathbf{B}_{j,\eta j+2}, \cdots, \mathbf{B}_{j,\eta j+17}$, respectively. Then

$$\mathbf{B}_{j1} \rightarrow \mathbf{B}(\mathbf{B}_j), \qquad \mathbf{B} \text{ conv } \mathbf{M}, \qquad \mathbf{B}_j \text{ conv } \mathbf{j}; \tag{1}$$

and, since $\mathbf{A}_{j1}, \cdots, \mathbf{A}_{j\eta j}, \mathfrak{A}_1, \cdots, \mathfrak{A}_{16}$ are provable,

$$\vdash \mathbf{B}_{ji} \qquad (i = 2, \cdots, \eta_j + 17). \tag{2}$$

Moreover, we may list all the other combinations occurring in the last mentioned derivation as $\mathbf{B}_{j,\eta j+18}, \cdots, \mathbf{B}_{j\theta_j}$ $(\theta_j > \eta_j + 17)$[14] in such a manner that

(3) for each i $(\eta_j + 18 \leqq i \leqq \theta_j)$, either there exist μ_{ji} and ν_{ji}, both $< i$, such that $\mathbf{B}_{ji}$ results from an application of R1 to $\mathbf{B}_{j\mu_{ji}}$, $\mathbf{B}_{j\nu_{ji}}$, or there exists $\mu_{ji} < i$ such that $\mathbf{B}_{ji}$ results from an application of a certain one of R2–R8 to $\mathbf{B}_{j\mu_{ji}}$; and

$$\mathbf{B}_{j\theta_j} \rightarrow \mathbf{N}', \qquad \mathbf{N}' \text{ conv } \mathbf{N}. \tag{4}$$

Let the terms of $\mathbf{B}$ be $\mathbf{b}_1, \cdots, \mathbf{b}_\tau$; of $\mathbf{B}_j$ be $\mathbf{b}_{j1}, \cdots, \mathbf{b}_{j\tau_j}$; of $\mathbf{B}_{ji}$ be $\mathbf{b}_{ji1}, \cdots, \mathbf{b}_{ji\tau_{ji}}$ $(i = 1, \cdots, \theta_j)$. Since $\mathbf{B}_j$ is $I(\mathbf{j}')$, $\tau_j > 1$. By 6III and (1), $\mathbf{b}_1, \cdots, \mathbf{b}_\tau, \mathbf{b}_{j1}, \cdots, \mathbf{b}_{j\tau_j}$ and $\mathbf{b}_{j11}, \cdots, \mathbf{b}_{j1\tau_{j1}}$ are the same series of terms, and $\tau + \tau_j = \tau_{j1}$. Let $\mathbf{B}'$, $\mathbf{B}'_j$, and $\mathbf{B}'_{ji}$ be the expressions obtained from $\mathbf{B}$, $\mathbf{B}_j$, and $\mathbf{B}_{ji}$, respectively, by the substitution of $t_1, \cdots, t_\tau; t_1, \cdots, t_{\tau_j}; t_1, \cdots, t_{\tau_{ji}}$ for the terms $\mathbf{b}_1, \cdots, \mathbf{b}_\tau; \mathbf{b}_{j1}, \cdots, \mathbf{b}_{j\tau_j}; \mathbf{b}_{ji1}, \cdots, \mathbf{b}_{ji\tau_{ji}}$, respectively. By 71, there exists a $\mathbf{G}$ such that $\mathbf{G}(1)$ conv $\lambda t_1 \cdots t_\tau \cdot \mathbf{B}'$ and $\mathbf{G}(2)$ conv λ & $et_1 \cdots t_\tau u \cdot e(I) \cdot e(t_1) \cdot \cdots \cdot e(t_\tau) \cdot u$. Similarly, there exists a $\mathbf{G}_j$ such that $\mathbf{G}_j(1)$ conv $\lambda t_1 \cdots t_{\tau_j} \cdot \mathbf{B}'_j$ and $\mathbf{G}_j(2)$ conv λ & $et_1 \cdots t_{\tau_j} \cdot e(t_1) \cdots \cdot e(t_{\tau_j})$ (since $\tau_j > 1$, & is a free symbol of $e(t_1) \cdots \cdot e(t_{\tau_j})$) ; and a $\mathbf{G}_{ji}$ $(i = 2, \cdots, \theta_j)$ such that $\mathbf{G}_{ji}(1)$ conv $\lambda t_1 \cdots t_{\tau_{ji}} \cdot \mathbf{B}'_{ji}$ and $\mathbf{G}_{ji}(2)$ conv λ & $et_1 \cdots t_{\tau_{ji}} \cdot e(I) \cdot e(t_1) \cdots \cdot e(t_{\tau_{ji}})$. Let $\mathbf{H} \rightarrow \lambda t_1 \cdots t_\tau x \cdot x(\mathbf{G}, t_1, \cdots, t_\tau)$, $\mathbf{H}_j \rightarrow \lambda t_1 \cdots t_{\tau_j} x \cdot x(\mathbf{G}_j, t_1, \cdots, t_{\tau_j})$, $\mathbf{H}_{j1} \rightarrow \lambda t_1 \cdots t_{\tau_{j1}} x \cdot x(\mathbf{G}, t_1, \cdots, t_\tau, x(\mathbf{G}_j, t_{\tau+1}, \cdots, t_{\tau_{j1}}))$, and $\mathbf{H}_{ji} \rightarrow \lambda t_1 \cdots t_{\tau_{ji}} x \cdot x(\mathbf{G}_{ji}, t_1, \cdots, t_{\tau_{ji}})$ $(i = 2, \cdots, \theta_j)$. Let $\mathbf{M}^* \rightarrow \mathbf{H}(\mathbf{b}_1, \cdots, \mathbf{b}_\tau)$, $\mathbf{j}^* \rightarrow \mathbf{H}_j(\mathbf{b}_{j1}, \cdots, \mathbf{b}_{j\tau_j})$, $\mathbf{B}^*_{ji} \rightarrow \mathbf{H}_{ji}(\mathbf{b}_{ji1}, \cdots, \mathbf{b}_{ji\tau_{ji}})$ $(i = 1, \cdots, \theta_j)$, $\mathbf{N}^* \rightarrow \mathbf{B}^*_{1\theta_1}$, $\mathfrak{N} \rightarrow \lambda y \cdot y(1)$, and $\mathfrak{L} \rightarrow \lambda y \cdot y(2, \&, E)$. For $i > 1$, our choice of $\underline{\mathbf{B}}^*_i$ was governed only by the form of $\mathbf{B}_{ji}$. According

[14] Our discussion does not apply verbatim to the case in which this derivation is of zero steps. But then we can use instead the derivation consisting of an application of R4 to $\mathbf{N}'$, followed by an application of R3 to the resulting expression $I(\mathbf{N}')$.

to (4), $\mathbf{B}_{1\theta_1}$ and $\mathbf{B}_{2\theta_2}$ are both the expression $\mathbf{N}'$. Hence we may suppose that $\mathbf{B}^*_{2\,\theta_2}$ was chosen to be the same expression as $\mathbf{B}^*_{1\,\theta_1}$. Then

(5) $\qquad \mathbf{B}^*_{j\,\theta_j}$ is $\mathbf{N}^*$.

$\mathbf{M}^*(\mathfrak{R})$ conv $\mathfrak{R}(\mathbf{G}, \mathbf{b}_1, \cdots, \mathbf{b}_\tau)$, conv $\mathbf{G}(1, \mathbf{b}_1, \cdots, \mathbf{b}_\tau)$, conv $\mathbf{B}$, conv $\mathbf{M}$. Thus

(6) $\qquad \mathbf{M}^*(\mathfrak{R})$ conv $\mathbf{M}$.

Similarly,

(7) $\qquad \mathbf{j}^*(\mathfrak{R})$ conv $\mathbf{j}$,

and $\mathbf{B}^*_{j\,i}(\mathfrak{R})$ conv $\mathbf{B}_{ji}$ $(i = 2, \cdots, \theta_j)$. Since both $\mathbf{B}^*_{j\,1}(\mathbf{X})$ and $\mathbf{M}^*(\mathbf{X}, \mathbf{j}^*(\mathbf{X}))$ are convertible into $\mathbf{X}(\mathbf{G}, \mathbf{b}_1, \cdots, \mathbf{b}_\tau, \mathbf{X}(\mathbf{G}_j, \mathbf{b}_{j1}, \cdots, \mathbf{b}_{j\tau_j}))$

(8) $\qquad$ for any $\mathbf{X}$, $\mathbf{B}^*_{j\,1}(\mathbf{X})$ conv $\mathbf{M}^*(\mathbf{X}, \mathbf{j}^*(\mathbf{X}))$.

Hence $\mathbf{B}^*_{j1}(\mathfrak{R})$ conv $\mathbf{M}^*(\mathfrak{R}, \mathbf{j}^*(\mathfrak{R}))$, conv $\mathbf{B}(\mathbf{B}_j)$, which is $\mathbf{B}_{j1}$. Thus, by cases $(i > 1, i = 1)$,

(9) $\qquad \mathbf{B}^*_{j\,i}(\mathfrak{R})$ conv $\mathbf{B}_{ji}$ $\qquad (i = 1, \cdots, \theta_j)$.

From (9) and (2) follows

(10) $\qquad \vdash \mathbf{B}^*_{j\,i}(\mathfrak{R}) \qquad (i = 2, \cdots, \eta_j + 17)$;

and from (9) and (4) follows

(11) $\qquad \mathbf{N}^*(\mathfrak{R})$ conv $\mathbf{N}$.

$\mathbf{B}^*_{j\,1}(\mathfrak{L})$ conv $\mathbf{G}(2, \&, E, \mathbf{b}_{j11}, \cdots, \mathbf{b}_{j1\tau}, \mathbf{G}_j(2, \&, E, \mathbf{b}_{j1,\tau+1}, \cdots, \mathbf{b}_{j1\tau_{j1}}))$, conv $E(I) \cdot E(\mathbf{b}_{j11}) \cdots \cdot E(\mathbf{b}_{j1\tau}) \cdot E(\mathbf{b}_{j1,\tau+1}) \cdots \cdot E(\mathbf{b}_{j1\tau_{j1}})$. Also, for $i = 2, \cdots, \theta_j$, $\mathbf{B}^*_{j\,i}(\mathfrak{L})$ conv $\mathbf{G}_{ji}(2, \&, E, \mathbf{b}_{ji1}, \cdots, \mathbf{b}_{ji\tau_{ji}})$, conv $E(I) \cdot E(\mathbf{b}_{ji1}) \cdots \cdot E(\mathbf{b}_{ji\tau_{ji}})$. Hence

(12) $\qquad \mathbf{B}^*_{j\,i}(\mathfrak{L})$ conv $E(I) \cdot E(\mathbf{b}_{ji1}) \cdots E(\mathbf{b}_{ji\tau_{ji}}) \qquad (i = 1, \cdots, \theta_j)$.

Any axiom $\mathbf{A}$ is convertible into $J(I, I, \mathbf{A}, I)$. Hence, by 5II, $\vdash E(I)$, and also, if $\mathbf{b}_{ji\tau}(1 \leqq r \leqq \tau_{ji})$ is an I or a J, then $\vdash E(\mathbf{b}_{ji\tau})$. If $\mathbf{b}_{ji\tau}$ is a constant symbol, then, by 5II and the definition of constant symbol, $E(\mathbf{b}_{ji\tau})$ can be inferred from some one of the axioms. It follows from 5VI, (1), and the definition of $\mathbf{j}$ $(j = 1, 2)$, that the free symbols of $\mathbf{B}_{j1}$ are the free symbols of $\mathbf{M}$. By 5V Corollary and (2), $\mathbf{B}_{ji}$ $(i = 2, \cdots, \eta_j + 17)$ contains no free variables. Since R1–R8 are rules the application of which introduces no new free symbols other than Σ, we conclude from (3) that the free variables of $\mathbf{B}_{ji}$ $(i = \eta_j + 18, \cdots, \theta_j)$ are included among the free variables of $\mathbf{B}_{j1}, \cdots, \mathbf{B}_{j,\,\eta_j+17}$, and hence among those of $\mathbf{M}$. Thus if the term $\mathbf{b}_{ji\tau}$ of the combination $\mathbf{B}_{ji}$ $(i = 1, \cdots, \theta_j)$ is a variable, it is a free variable of $\mathbf{M}$. In this case, by 5II, $E(\mathbf{M}) \vdash E(\mathbf{b}_{ji\tau})$. But $\mathbf{b}_{ji\tau}$, since a term of a combination, is either an I, a J, a constant symbol, or a variable. Hence, if $E(\mathbf{M})$ is assumed, we can prove $E(\mathbf{b}_{ji\tau})$ $(r = 1, \cdots, \tau_{ji})$, and also $E(I)$. It follows, using (12), that

(13) $$E(\mathbf{M}) \vdash \mathbf{B}^{*}_{j\,i}(\mathfrak{L}) \qquad (i = 1, \cdots, \theta_j).$$

Supposing now that $\eta_j + 18 \leqq i \leqq \theta_j$, we distinguish cases in accordance with the possibilities under (3).

Case 1: $\mathbf{B}_{j\,i}$ results from an application of R1 to $\mathbf{B}_{j\mu_{ji}}$, $\mathbf{B}_{j\nu_{ji}}$. Then $\mathbf{B}_{j\mu_{ji}}$, $\mathbf{B}_{j\nu_{ji}}$, $\mathbf{B}_{j\,i}$ must be of the forms $\Pi(\mathbf{f}^{ji}, \mathbf{g}^{ji})$, $\mathbf{f}^{ji}(\mathbf{p}^{ji})$, $\mathbf{g}^{ji}(\mathbf{p}^{ji})$, respectively, where by 6II the parts $\mathbf{f}^{ji}$, $\mathbf{g}^{ji}$, $\mathbf{p}^{ji}$ are combinations. Let the sequence of terms of $\mathbf{f}^{ji}$ be $\mathbf{f}^{ji}_1, \cdots, \mathbf{f}^{ji}_{\alpha_{ji}}$; of $\mathbf{g}^{ji}$ be $\mathbf{g}^{ji}_1, \cdots, \mathbf{g}^{ji}_{\beta_{ji}}$; and of $\mathbf{p}^{ji}$ be $\mathbf{p}^{ji}_1, \cdots, \mathbf{p}^{ji}_{\gamma_{ji}}$. By 6III, (i) II, $\mathbf{f}^{ji}_1, \cdots, \mathbf{f}^{ji}_{\alpha_{ji}}, \mathbf{g}^{ji}_1, \cdots, \mathbf{g}^{ji}_{\beta_{ji}}, \mathbf{f}^{ji}_1, \cdots, \mathbf{f}^{ji}_{\alpha_{ji}}, \mathbf{p}^{ji}_1, \cdots, \mathbf{p}^{ji}_{\gamma_{ji}}; \mathbf{g}^{ji}_1, \cdots, \mathbf{g}^{ji}_{\beta_{ji}}, \mathbf{p}^{ji}_1, \cdots, \mathbf{p}^{ji}_{\gamma_{ji}}$ and $\mathbf{b}_{j\mu_{ji}1}, \cdots, \mathbf{b}_{j\nu_{ji}\tau_{j\mu_{ji}}}; \mathbf{b}_{j\nu_{ji}1}, \cdots, \mathbf{b}_{j\nu_{ji}\tau_{j\nu_{ji}}}; \mathbf{b}_{ji1}, \cdots, \mathbf{b}_{ji\tau_{ji}}$, respectively, are the same series of expressions. Let $\mathbf{f}_1, \cdots, \mathbf{f}_{\alpha_{ji}}, \mathbf{g}_1, \cdots, \mathbf{g}_{\beta_{ji}}, \mathbf{p}_1, \cdots, \mathbf{p}_{\gamma_{ji}}$ be a set of variables distinct from each other and from $\mathbf{x}$. Let $\bar{\mathbf{f}}^{ji}$ be the expression obtained from $\mathbf{f}^{ji}$ by the substitution of $\mathbf{f}_1, \cdots, \mathbf{f}_{\alpha_{ji}}$ for the terms $\mathbf{f}^{ji}_1, \cdots, \mathbf{f}^{ji}_{\alpha_{ji}}$, respectively; and define similarly $\bar{\mathbf{g}}^{ji}$ and $\bar{\mathbf{p}}^{ji}$. Let $\bar{\mathbf{B}}^{i}_{j\mu_{ji}}, \bar{\mathbf{B}}^{i}_{j\nu_{ji}}, \bar{\mathbf{B}}_{ji}$ be the expressions $\Pi(\bar{\mathbf{f}}^{ji}, \bar{\mathbf{g}}^{ji}), \bar{\mathbf{f}}^{ji}(\bar{\mathbf{p}}^{ji})$, $\bar{\mathbf{g}}^{ji}(\bar{\mathbf{p}}^{ji})$, respectively.

By Rule V, $\bar{\mathbf{B}}^{i}_{j\mu_{ji}}, \bar{\mathbf{B}}^{i}_{j\nu_{ji}} \vdash \bar{\mathbf{B}}_{ji}$. $[\Sigma(x) \supset {}_x\Sigma(x)] \cdot \{\lambda x \cdot \Sigma(x)\}(\Sigma)$ is an example of a provable formula of the form $\Pi(\mathbf{f}, \mathbf{g}) \cdot \mathbf{f}(\mathbf{p})$. (It is provable, since any axiom $\mathbf{A}$ is convertible into $I(\mathbf{A})$. Thence, by an application of Rule IV, $\Sigma(I)$; by a second application of Rule IV, $\Sigma(\Sigma)$; and by conversion, $\{\lambda x \cdot \Sigma(x)\}(\Sigma)$. Thence $\Sigma x \cdot \Sigma(x)$ is provable by IV. Also $\Sigma(x) \vdash \Sigma(x)$. Hence, by Theorem I, $\Sigma(x) \supset {}_x\Sigma(x)$.) By 6IV, $[\Sigma(x) \supset {}_x\Sigma(x)] \cdot \{\lambda x \cdot \Sigma(x)\}(\Sigma)$ is convertible into $S^{\mathbf{f}_1 \cdots \mathbf{f}_{\alpha_{ji}-1}\mathbf{f}_{\alpha_{ji}}}_{I \cdots I \quad \lambda x \cdot \Sigma(x)\ I \cdots I}\ S^{\mathbf{g}_1 \cdots \mathbf{g}_{\beta_{ji}-1}\mathbf{g}_{\beta_{ji}}}_{I \cdots I \quad \lambda x \cdot \Sigma(x)\ I \cdots I}\ S^{\mathbf{p}_1 \cdots \mathbf{p}_{\gamma_{ji}-1}\mathbf{p}_{\gamma_{ji}}}_{I \cdots I \quad \Sigma}\ \bar{\mathbf{B}}^{i}_{j\mu_{ji}} \cdot \bar{\mathbf{B}}^{i}_{j\nu_{ji}} \Big|$, which yields, by alternate applications of III and IV, $\Sigma \mathbf{f}_1 \cdots \mathbf{f}_{\alpha_{ji}} \mathbf{g}_1 \cdots \mathbf{g}_{\beta_{ji}} \mathbf{p}_1 \cdots \mathbf{p}_{\gamma_{ji}} \cdot \bar{\mathbf{B}}^{i}_{j\mu_{ji}} \cdot \bar{\mathbf{B}}^{i}_{j\nu_{ji}}$. Hence, by a corollary of Theorem I, (ii)

$$[\bar{\mathbf{B}}^{i}_{j\mu_{ji}} \cdot \bar{\mathbf{B}}^{i}_{j\nu_{ji}}] \supset {}_{\mathbf{f}1 \cdots \mathbf{f}_{\alpha_{ji}}\mathbf{g}_1 \cdots \mathbf{g}_{\beta_{ji}}\mathbf{p}_1 \cdots \mathbf{p}_{\gamma_{ji}}} \cdot \bar{\mathbf{B}}_{ji}.$$

Now let

$$\bar{\mathbf{B}}^{i*}_{j\mu_{ji}} \to \mathbf{H}_{j\mu_{ji}}(\Pi, \mathbf{f}_1, \cdots, \mathbf{f}_{\alpha_{ji}}, \mathbf{g}_1, \cdots \mathbf{g}_{\beta_{ji}}), \quad \bar{\mathbf{B}}^{i*}_{j\nu_{ji}} \to \mathbf{H}_{j\nu_{ji}}(\mathbf{f}_1, \cdots, \mathbf{f}_{\alpha_{ji}}, \mathbf{p}_1, \cdots, \mathbf{p}_{\gamma_{ji}}),$$

$$\bar{\mathbf{B}}^{*}_{j\,i} \to \mathbf{H}_{j\,i}(\mathbf{g}_1, \cdots, \mathbf{g}_{\beta_{ji}}, \mathbf{p}_1, \cdots, \mathbf{p}_{\gamma_{ji}}),$$

and

$$\mathbf{C}_{j\,i} \to \lambda \mathbf{x} \cdot [\bar{\mathbf{B}}^{i*}_{j\mu_{ji}}(\mathbf{x}) \cdot \bar{\mathbf{B}}^{i*}_{j\nu_{ji}}(\mathbf{x})] \supset {}_{\mathbf{f}_1 \cdots \mathbf{f}_{\alpha_{ji}}\mathbf{g}_1 \cdots \mathbf{g}_{\beta_{ji}}\mathbf{p}_1 \cdots \mathbf{p}_{\gamma_{ji}}} \cdot \bar{\mathbf{B}}^{*}_{j\,i}(x).$$

Note that $\mathbf{H}_{j\,s}(s = \mu_{j\,i}, \nu_{j\,i})$, $\mathbf{H}_{j\,i}$ contain no free symbols.

Similarly to (9), $\bar{\mathbf{B}}^{i*}_{j\,s}(\mathfrak{L})$ conv $\bar{\mathbf{B}}^{i}_{j\,s}$ and $\bar{\mathbf{B}}^{*}_{j\,i}(\mathfrak{L})$ conv $\bar{\mathbf{B}}_{j\,i}$. Hence from (ii) follows (iii) $\vdash \mathbf{C}_{ji}(\mathfrak{K})$.

Similarly to (12), we have $\bar{\mathbf{B}}^{i*}_{j\mu_{ji}}(\mathfrak{L}) \cdot \bar{\mathbf{B}}^{i*}_{j\nu_{ji}}(\mathfrak{L})$ conv $[E(I) \cdot E(\Pi) \cdot E(\mathbf{f}_1) \cdot \cdots \cdot E(\mathbf{f}_{\alpha_{ji}}) \cdot E(\mathbf{g}_1) \cdot \cdots \cdot E(\mathbf{g}_{\beta_{ji}})] \cdot E(I) \cdot E(\mathbf{f}_1) \cdot \cdots \cdot E(\mathbf{f}_{\alpha_{ji}}) \cdot E(\mathbf{p}_1) \cdot \cdots \cdot E(\mathbf{p}_{\gamma_{ji}})$, and $\bar{\mathbf{B}}^{*}_{j\,i}(\mathfrak{L})$ conv $E(I) \cdot E(\mathbf{g}_1) \cdot \cdots \cdot E(\mathbf{g}_{\alpha_{ji}}) \cdot E(\mathbf{p}_1) \cdot \cdots \cdot E(\mathbf{p}_{\gamma_{ji}})$. Since the factors of the latter logical product are included among those of the former, $\bar{\mathbf{B}}^{i*}_{j\mu_{ji}}(\mathfrak{L}) \cdot \bar{\mathbf{B}}^{i*}_{j\nu_{ji}}(\mathfrak{L}) \vdash \bar{\mathbf{B}}^{*}_{j\,i}(\mathfrak{L})$. $[E(I) \cdot E(\Pi) \cdot E(I) \cdot \cdots \cdot E(I)] \cdot E(I) \cdot \cdots \cdot E(I)$ $(2 + \tau_{j\mu_{ji}} + \tau_{j\nu_{ji}}$ factors) is provable,

and convertible into $\{\lambda \mathbf{f}_1 \cdots \mathbf{p}_{\gamma_{ji}} \cdot \overline{\mathbf{B}}^{i*}_{j\mu_{ji}}(\mathfrak{Q}) \cdot \overline{\mathbf{B}}^{i*}_{j\nu_{ji}}(\mathfrak{Q})\}(I, \cdots, I)$, which by 5III yields $\Sigma \mathbf{f}_1 \cdots \mathbf{p}_{\gamma_{ji}} \cdot \overline{\mathbf{B}}^{i*}_{j\mu_{ji}}(\mathfrak{Q}) \cdot \overline{\mathbf{B}}^{i*}_{j\nu_{ji}}(\mathfrak{Q})$. Hence, by a corollary of Theorem I, (iv) $\vdash \mathbf{C}_{ji}(\mathfrak{Q})$.

It follows from (i) by applications of 5I that $\mathbf{B}^*_{j\mu_{ji}}(\mathbf{x}) \cdot \mathbf{B}^*_{j\nu_{ji}}(\mathbf{x})$ conv $\{\lambda \mathbf{f}_1 \cdots \mathbf{p}_{\gamma_{ji}} \cdot \overline{\mathbf{B}}^{i*}_{j\mu_{ji}}(\mathbf{x}) \cdot \overline{\mathbf{B}}^{i*}_{j\nu_{ji}}(\mathbf{x})\}(\mathbf{f}_1^{ji}, \cdots, \mathbf{p}^{ji}_{\gamma_{ji}})$ and $\mathbf{B}^*_{ji}(\mathbf{x})$ conv $\{\lambda \mathbf{g}_1 \cdots \mathbf{p}_{\gamma_{ji}} \cdot \overline{\mathbf{B}}^{i*}_{ji}(\mathbf{x})\}$ $(\mathbf{g}_1^{ji}, \cdots, \mathbf{p}^{ji}_{\gamma_{ji}})$. Hence, by 5IV, (v) $\mathbf{C}_{ji}(\mathbf{x}), \mathbf{B}^*_{j\mu_{ji}}(\mathbf{x}), \mathbf{B}^*_{j\nu_{ji}}(\mathbf{x}) \vdash \mathbf{B}^*_{ji}(\mathbf{x})$.

Cases 2–8: $\mathbf{B}_{ji}$ results from an application of R2–R8, respectively, to $\mathbf{B}_{j\mu_{ji}}$. For each of these cases, a function $\mathbf{C}_{ji}$ can be defined and shown to have the properties (iii), (iv), and (v′) $\mathbf{C}_{ji}(\mathbf{x}), \mathbf{B}^*_{j\mu_{ji}}(\mathbf{x}) \vdash \mathbf{B}^*_{ji}(\mathbf{x})$ in the same manner as in Case 1. $\overline{\mathbf{B}}^{i}_{j\mu_{ji}} \vdash \overline{\mathbf{B}}_{ji}$ in Case 2 by application of Rule IV, and in Cases 3–8 by conversions, utilizing the definitions of I and J. Examples of provable formulas of the forms occurring in the hypotheses of R2–R8 are $\Sigma(\Sigma)$, $I(\Sigma(\Sigma))$, $\Sigma(\Sigma)$, $I(I(\Sigma, \Sigma))$, $I(\Sigma(\Sigma))$, $I(J(I, I, \Sigma, \Sigma))$, $I(I(I, I(\Sigma, \Sigma)))$, respectively. In Case 2, $\mathbf{C}_{ji} \rightarrow \lambda \mathbf{x} \cdot \overline{\mathbf{B}}^{i*}_{j\mu_{ji}}(\mathbf{x}) \supset_{\mathbf{p}_1 \cdots \mathbf{p}_{\gamma_{ji}} \mathbf{f}_1 \cdots \mathbf{f}_{\alpha_{ji}}} \cdot \overline{\mathbf{B}}^{i*}_{ji}(\mathbf{x})$.

Thus it is shown by cases that there exist functions $\mathbf{C}_{ji}$ such that

(14) $\quad \vdash \mathbf{C}_{ji}(\mathfrak{R}) \qquad (i = \eta_j + 18, \cdots, \theta_j)$,

(15) $\quad \vdash \mathbf{C}_{ji}(\mathfrak{Q}) \qquad$ " " ,

(16) $\mathbf{C}_{ji}(\mathbf{x}), \mathbf{B}^*_{j\mu_{ji}}(\mathbf{x})(, \mathbf{B}^*_{j\nu_{ji}}(\mathbf{x})) \vdash \mathbf{B}_{ji}(\mathbf{x})$ for a certain $\mu_{ji}(, \nu_{ji}) < i$ " " .

Expressions $\mathbf{M}^*$, $\mathbf{j}^*$, $\mathbf{N}^*$, $\mathfrak{R}$ have been defined and shown to possess the properties (6), (7), (11). To complete the proof of 8I, we must define $\mathbf{T}$ and show that the conditions $\vdash \mathbf{T}(\mathfrak{R})$ and $E(\mathbf{M}) \vdash \mathbf{M}^*(\mathbf{x}, \mathbf{j}^*(\mathbf{x}))\mathbf{T}(\mathbf{x}) \supset_{\mathbf{x}} \mathbf{N}^*(\mathbf{x})$ are satisfied. Let $\mathbf{T}_j \rightarrow \lambda \mathbf{x} \cdot \mathbf{B}^*_{j2}(\mathbf{x}) \cdots \mathbf{B}^*_{j,\,\eta_j+17}(\mathbf{x}) \mathbf{C}_{j,\,\eta_j+18}(\mathbf{x}) \cdots \mathbf{C}_{j\theta_j}(\mathbf{x})$, and $\mathbf{T} \rightarrow \lambda \mathbf{x} \cdot \mathbf{T}_1(\mathbf{x}) \mathbf{T}_2(\mathbf{x})$. It follows from the definition of $\mathbf{B}^*_{ji}$ $(i = 1, \cdots, \theta_j)$, and the previously noted fact that the only variables among the $\mathbf{b}_{jir}$ are free variables of $\mathbf{M}$, that $\mathbf{x}$ does not occur in $\mathbf{B}^*_{ji}$ as a free variable. $\mathbf{C}_{ji}$ $(i = \eta_j + 18, \cdots, \theta_j)$ contains no free variables. Hence, for any $\mathbf{X}$, $\mathbf{T}_j(\mathbf{X})$ conv $\mathbf{B}^*_{j2}(\mathbf{X}) \cdots \mathbf{C}_{j\theta_j}(\mathbf{X})$, $\mathbf{T}(\mathbf{X})$ conv $\mathbf{T}_1(\mathbf{X})\mathbf{T}_2(\mathbf{X})$, and $\mathbf{B}^*_{j1}(\mathbf{X})\mathbf{T}(\mathbf{X})$ conv $\{\lambda \mathbf{x} \cdot \mathbf{B}^*_{j1}(\mathbf{x})\mathbf{T}(\mathbf{x})\}(\mathbf{X})$. Using the first two of these conversions, $\vdash \mathbf{T}(\mathfrak{R})$ follows from (10) and (14).

Assume $E(\mathbf{M})$. According to (13), $\mathbf{B}^*_{j1}(\mathfrak{Q}), \cdots, \mathbf{B}^*_{j,\,\eta_j+17}(\mathfrak{Q})$ are provable as consequences; and, according to (15), $\mathbf{C}_{j,\,\eta_j+18}(\mathfrak{Q}), \cdots, \mathbf{C}_{j\theta_j}(\mathfrak{Q})$ are provable. We can infer $\mathbf{B}^*_{j1}(\mathfrak{Q})\mathbf{T}(\mathfrak{Q})$, and thence, by conversion and Rule IV, $\Sigma \mathbf{x} \cdot \mathbf{B}^*_{j1}(\mathbf{x})\mathbf{T}(\mathbf{x})$.

Assume $\mathbf{B}^*_{j1}(\mathbf{x})\mathbf{T}(\mathbf{x})$. Thence we can infer $\mathbf{B}^*_{j1}(\mathbf{x}), \cdots, \mathbf{B}^*_{j,\,\eta_j+17}(\mathbf{x}), \mathbf{C}_{j,\,\eta_j+18}(\mathbf{x}), \cdots, \mathbf{C}_{j\theta_j}(\mathbf{x})$. According to (16), we can prove $\mathbf{B}^*_{j,\,\eta_j+18}(\mathbf{x})$ from $\mathbf{C}_{j,\,\eta_j+18}(\mathbf{x})$ and a certain one or two of $\mathbf{B}^*_{j1}(\mathbf{x}), \cdots, \mathbf{B}^*_{j,\,\eta_j+17}(\mathbf{x})$; then (if $\eta_j + 18 < \theta_j$) prove $\mathbf{B}^*_{j,\,\eta_j+19}(\mathbf{x})$ from $\mathbf{C}_{j,\,\eta_j+19}(\mathbf{x})$ and a certain one or two of $\mathbf{B}^*_{j1}(\mathbf{x}), \cdots, \mathbf{B}^*_{j,\,\eta_j+18}(\mathbf{x})$; and so on. Thus $\mathbf{B}^*_{j,\,\eta_j+18}(\mathbf{x}), \cdots, \mathbf{B}^*_{j\theta_j}(\mathbf{x})$ are provable successively as consequences of $\mathbf{B}^*_{j1}(\mathbf{x})\mathbf{T}(\mathbf{x})$.

We have now shown that $E(\mathbf{M}) \vdash \Sigma \mathbf{x} \cdot \mathbf{B}^*_{j1}(\mathbf{x})\mathbf{T}(\mathbf{x})$, and $\mathbf{B}^*_{j1}(\mathbf{x})\mathbf{T}(\mathbf{x}) \vdash \mathbf{B}^*_{j\theta_j}(\mathbf{x})$.

Hence, by Theorem I (i.e., by Theorem I with "$\vdash$" replaced by "$E(\mathbf{M}) \vdash$," and "variable" by "variable distinct from the free variables of $E(\mathbf{M})$"),[15] $E(\mathbf{M}) \vdash \mathbf{B}^*_{j1}(\mathbf{x})\mathbf{T}(\mathbf{x}) \supset_{\mathbf{x}} \mathbf{B}^*_{j\theta_j}(\mathbf{x})$. Using (8) and (5), it follows that

$$E(\mathbf{M}) \vdash \mathbf{M}^*(\mathbf{x}, \mathbf{j}^*(\mathbf{x}))\mathbf{T}(\mathbf{x}) \supset_{\mathbf{x}} \mathbf{N}^*(\mathbf{x}).$$

9. **Proof by cases.**

9I. *If* $\mathbf{B} = 1 \vdash \mathbf{C}$ *and* $\mathbf{B} = 2 \vdash \mathbf{C}$, *then* $M(\mathbf{B}) \vdash \mathbf{C}$.

Let $\mathbf{x}$ be a variable distinct from the free variables of $\mathbf{B}$. Then from the hypothesis it follows that $\{\lambda\mathbf{x}\cdot\mathbf{B} = \mathbf{x}\}(\mathbf{j}) \vdash \mathbf{C}$ $(j = 1, 2)$.[16] Hence, according to 8I, there exist expressions $\mathbf{F}, \mathbf{j}^*, \mathbf{T}, \mathbf{C}^*, \mathfrak{R}$ such that $\mathbf{F}(\mathfrak{R})$ conv $\lambda\mathbf{x}\cdot\mathbf{B} = \mathbf{x}, \mathbf{j}^*(\mathfrak{R})$ conv $\mathbf{j}, \vdash \mathbf{T}(\mathfrak{R}), \mathbf{C}^*(\mathfrak{R})$ conv $\mathbf{C}$, and $E(\lambda\mathbf{x}\cdot\mathbf{B} = \mathbf{x}) \vdash \mathbf{F}(\mathbf{x}, \mathbf{j}^*(\mathbf{x}))\mathbf{T}(\mathbf{x}) \supset_{\mathbf{x}} \mathbf{C}^*(\mathbf{x})$.

From 5VI, the fact $\mathbf{j}^*(\mathfrak{R})$ conv $\mathbf{j}$, and the definition of $\mathbf{j}$, it follows that $\mathbf{j}^*(\mathfrak{R})$, and hence $\mathbf{j}^*$, contains no free symbols. Hence, by 7I, there exists a $\mathbf{J}$ such that $\mathbf{J}(\mathbf{j})$ conv $\mathbf{j}^*$.

Assume $M(\mathbf{B})$. Then, by 5II, $E(\mathbf{B})$. From this formula and Theorem 1, Church, 1933, p. 844, we obtain, by conversion and Rule V, the formula $\mathbf{B} = \mathbf{B}$. Thence, by conversion, $\{\lambda\mathbf{x}\cdot\mathbf{B} = \mathbf{x}\}$ $(\mathbf{B})$; and, by 5II, $E(\lambda\mathbf{x}\cdot\mathbf{B} = \mathbf{x})$. Hence $\mathbf{F}(\mathbf{x}, \mathbf{j}^*(\mathbf{x}))\mathbf{T}(\mathbf{x}) \supset_{\mathbf{x}} \mathbf{C}^*(\mathbf{x})$; and by conversion, (1) $\mathbf{F}(\mathbf{x}, \mathbf{J}(\mathbf{j}, \mathbf{x}))\mathbf{T}(\mathbf{x}) \supset_{\mathbf{x}} \mathbf{C}^*(\mathbf{x})$. $M(\mathbf{B})$ is convertible into $\mathbf{x}(1)\mathbf{x}(2) \supset_{\mathbf{x}} \mathbf{x}(\mathbf{B})$. From this implication, and the logical product of the two formulas (1), we infer (2) $\mathbf{F}(\mathbf{x}, \mathbf{J}(\mathbf{B}, \mathbf{x}))\mathbf{T}(\mathbf{x}) \supset_{\mathbf{x}} \mathbf{C}^*(\mathbf{x})$ by Rule V and conversion.

In view of 7.1, 7.2, 5II, and Theorem 1, $\mathbf{j} = \mathbf{j}$ is provable. $\mathbf{j} = \mathbf{j}$ conv $\mathbf{j} = \mathbf{j}^*(\mathfrak{R})$, conv $\mathbf{j} = \mathbf{J}(\mathbf{j}, \mathfrak{R})$. Hence $1 = \mathbf{J}(1, \mathfrak{R})\cdot 2 = \mathbf{J}(2, \mathfrak{R})$ is provable; and, using $M(\mathbf{B})$, we can infer $\mathbf{B} = \mathbf{J}(\mathbf{B}, \mathfrak{R})$ by Rule V and conversion. $\mathbf{B} = \mathbf{J}(\mathbf{B}, \mathfrak{R})$ conv $\{\lambda\mathbf{x}\cdot\mathbf{B} = \mathbf{x}\}(\mathbf{J}(\mathbf{B}, \mathfrak{R}))$, conv $\mathbf{F}(\mathfrak{R}, \mathbf{J}(\mathbf{B}, \mathfrak{R}))$. From (2), and the logical product of $\mathbf{F}(\mathfrak{R}, \mathbf{J}(\mathbf{B}, \mathfrak{R}))$ and the provable formula $\mathbf{T}(\mathfrak{R})$, we infer $\mathbf{C}^*(\mathfrak{R})$ by V and conversion. $\mathbf{C}^*(\mathfrak{R})$ conv $\mathbf{C}$. Thus, under the hypothesis of the theorem, $\mathbf{C}$ is shown to be provable from $M(\mathbf{B})$.

10. **Reductio ad absurdum.**

10I. *If the free variables of* $\mathbf{C}$ *are* $\mathbf{y}_1, \cdots, \mathbf{y}_m$ $(m \geqq 0)$, *then* $1 = 2$, $E(\mathbf{y}_1) \cdots E(\mathbf{y}_m) \vdash \mathbf{C}$.

If the free symbols of $\mathbf{C}$ are $\mathbf{z}_1, \cdots, \mathbf{z}_n$, by 7I there is an $\mathbf{L}$ such that $\mathbf{L}(1)$ conv $E(\&)E(\mathbf{z}_1) \cdots E(\mathbf{z}_n)$ and $\mathbf{L}(2)$ conv $\mathbf{C}\cdot E(\&)$. Then $E(\mathbf{y}_1) \cdots E(\mathbf{y}_m) \vdash \mathbf{L}(1)$. $1 = 2$ conv $\phi(1) \supset_{\phi} \phi(2)$, which with $\mathbf{L}(1)$ yields $\mathbf{L}(2)$. $\mathbf{L}(2) \vdash \mathbf{C}$.

10II. *If* $\mathbf{B} = 1 \vdash \mathbf{C}$ *and* $\mathbf{B} = 2 \vdash \mathbf{D} = 1$, *then* $M(\mathbf{B})$, $\mathbf{D} = 2 \vdash \mathbf{C}$. (*Similarly, interchanging* $\mathbf{B} = 1$ *and* $\mathbf{B} = 2$, *or* (*and*) $\mathbf{D} = 1$ *and* $\mathbf{D} = 2$.)

Since $\mathbf{B} = 1 \vdash \mathbf{C}$, the free variables, $\mathbf{y}_1, \cdots, \mathbf{y}_m$, of $\mathbf{C}$ are included among those of $\mathbf{B} = 1$ (by 5V), and hence among those of $\mathbf{B} = 2$. Hence, by 5II, $\mathbf{B} = 2 \vdash E(\mathbf{y}_1) \cdots E(\mathbf{y}_m)$. By hypothesis, $\mathbf{B} = 2 \vdash \mathbf{D} = 1$. In view of

[15] Cf. §4. The rule stated in §4 will sometimes be used tacitly.

[16] In the present proof we employ j under the same conventions as in §8.

Theorems 2, 3, Church, 1933, p. 845, $\mathbf{D} = 1, \mathbf{D} = 2 \vdash 1 = 2$. Hence $\mathbf{B} = 2, \mathbf{D} = 2 \vdash 1 = 2 \cdot E(\mathbf{y}_1) \cdots E(\mathbf{y}_m)$. It follows by 10I that $\mathbf{B} = 2, \mathbf{D} = 2 \vdash \mathbf{C}$. Now, by 9I (with "$\vdash$" replaced by "$\mathbf{D} = 2 \vdash$"), we may conclude that $M(\mathbf{B}), \mathbf{D} = 2 \vdash \mathbf{C}$.

PRINCETON UNIVERSITY, PRINCETON, N. J.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

THE CRITICAL POINT THEORY UNDER GENERAL BOUNDARY CONDITIONS

BY MARSTON MORSE AND GEORGE BOOTH VAN SCHAACK

(Received February 2, 1934)

1. **Introduction.** The present paper has two purposes. The first is to simplify as far as possible the treatment of the critical points of non-degenerate functions of n real variables. The second is to extend the earlier treatment of Morse to the case of general boundary conditions.

A method has been devised by means of which the given function f can be replaced without loss of generality by a new function φ whose Taylor's expansion neighboring each critical point ends with terms of the second order. The orthogonal trajectories to the contour manifolds of f neighboring the critical points would in general require a somewhat difficult analysis. But in the case of the function φ the major part of these difficulties disappears.

The sections on the critical points of f under general boundary conditions are entirely new. Theorem 10, however, was stated in the reference Morse (10) given at the end of the paper. It is hoped that the present paper may serve as an introduction to the theory of critical points of a function of n variables. For that reason the difficult case of degenerate critical loci has not been taken up. Different methods would be required for the analysis of the degenerate case. See, for example, Morse (12), Chapter VI. The general degenerate case under general boundary conditions has been treated by the authors and the results will be published shortly.

FORMULATION OF THE PROBLEM

2. **The region R and the function f.** Let R be an open, n-dimensional region, lying in a bounded part of a euclidean space of n-dimensions of coordinates $(x_1, x_2, \cdots, x_n)$. In R let there be defined a real, single-valued function $f(x_1, x_2, \cdots, x_n)$ of class[1] C''. A point at which all of the first partial derivatives of f vanish will be called a *critical point* of f, and the value of f at such a point, a *critical value* of f. All other points of R will be called *ordinary points* of f. We shall assume that the critical points of f are *non-degenerate*, that is, that the determinant

$$|f_{x_i x_j}| \neq 0 \qquad (i, j = 1, \cdots, n) \tag{2.1}$$

at a critical point. A function all of whose critical points are non-degenerate

[1] A function is said to be of class $C^{(m)}$ if it is continuous together with all of its derivatives up to and including those of the m^{th} order.

will be called a *non-degenerate function*. It follows from condition (2.1) that the critical points of f are isolated. For by the theory of implicit functions, the equations $f_{x_i} = 0$, $(i = 1, \cdots, n)$ will then have in the neighborhood of any solution only that solution.

Moreover, in any closed region S all of whose points are points of R, there is at most a finite number of critical points of f. Suppose there were an infinite set of such points in S. This set would have at least one cluster point, P. Since S is closed P would be a point of S. Since the derivatives f are continuous and since they all vanish on the set in question, they would all vanish at P. Thus P would be a non-isolated critical point of f, contrary to the result obtained above.

Where there is no ambiguity we shall write $f(x_1, \cdots, x_n)$ as $f(x)$. Throughout the paper we shall use the convention that the repetition of a subscript indicates summation with respect to that subscript. With this understood we now define the index of a critical point. Let (x^0) be a critical point of f. For convenience let (x^0) be the origin. Suppose $f(0) = 0$. Then there exists a real, non-singular, linear, homogeneous transformation of the variables (x) into variables (y) under which

$$f_{x_i x_j}(0)\, x_i\, x_j \equiv -y_1^2 - \cdots - y_k^2 + y_{k+1}^2 + \cdots + y_n^2,$$

where k is one of the integers $0, 1, \cdots, n$. The integer k is called the *index* of the critical point. It is independent of the particular linear transformation by means of which the above reduction is made. At a critical point of index zero f has a relative minimum, and at one of index n a relative maximum.

3. **The boundary conditions α.** Before proceeding further it is convenient to define the *boundary conditions* α. These conditions are as follows:

Iα. *There shall exist a closed subdomain S of R which contains all the critical points of f in its interior.*

IIα. *The boundary A of S shall be identical with a locus $f = M$, a constant.*

IIIα. *The value M of f on A shall be greater than the value of f at any interior point of S.*

We shall refer to these conditions more explicitly as the boundary conditions α on f *relative* to S.

4. **The complex of points satisfying $f \leqq c$.** Let c be any ordinary value of f at most equal to the absolute maximum M of f in S. Then, as Cairns (4) has shown, the set of points of S satisfying the condition $f \leqq c$ is an n-dimensional complex in the technical sense of analysis situs. Its boundary consists of the points of S which satisfy $f = c$.

We shall denote the connectivity numbers, modulo 2, of an n-dimensional complex by $R_0, R_1, \cdots, R_n$. We take these numbers in the sense of Alexander (1) and Lefschetz (8). The numbers $R_1, \cdots, R_n$ as defined by Veblen (15) are one larger than those used by Alexander and Lefschetz. See also Tucker (14).

It is the purpose of the first part of this paper to derive the relations which

exist between the numbers, M_k, of the critical points of index k of f in S and the connectivity numbers of the complex S.

5. **Related functions.** The relations which we shall derive depend merely upon the connectivities of S together with the critical points of f and their indices. Two functions f and φ which are of class C'' in R, whose critical points are non-degenerate and are the same in position and index will be termed *related*. If two functions f and φ are related and both satisfy the boundary conditions α relative to the same region S the relations between their critical points and the connectivities of S will be the same.

We are led to the following lemma.

LEMMA 1. *There exists a function φ which is related to f, which satisfies the boundary conditions α relative to the domain S, whose Taylor's expansion neighboring each critical point consists merely of a constant plus a quadratic form, and which takes on distinct critical values at distinct critical points.*

Let (x^0) be an arbitrary critical point of f. For simplicity we take (x^0) as the origin. Suppose $f(0) = 0$. Let $h(z)$ be any function of class C'', $0 \leqq z$, such that

$$\begin{aligned} h(z) &\equiv 1, & 0 \leqq z \leqq 1, \\ 0 \leqq h(z) &\leqq 1, & 1 < z < 4, \\ h(z) &\equiv 0, & 4 \leqq z. \end{aligned}$$

Let r^2 be the sum $x_i x_i$. Let e be a positive constant. We shall suitably restrict e in size during the course of the proof.

We define a function $\bar{\varphi}(x)$ by the following identity:

$$\bar{\varphi}(x) \equiv f(x) + h\left(\frac{r^2}{e^2}\right)\left[\frac{\alpha_{ij}\, x_i\, x_j}{2} - f(x)\right],$$

where

$$\alpha_{ij} = f_{x_i x_j}(0) \qquad (i, j = 1, \cdots, n).$$

The function $\bar{\varphi}$ is defined in R.

We seek the critical points of $\bar{\varphi}$. We take e so small that the origin is the only critical point of f for which $r \leqq 2e$. Now

$$\bar{\varphi} \equiv f, \qquad r \geqq 2e.$$

Hence on and outside of the sphere of radius $2e$ with center at the origin $\bar{\varphi}$ has critical points which are the same in position and index as those of f. On and within the sphere of radius e with center at the origin we have

$$\bar{\varphi} \equiv \frac{\alpha_{ij}\, x_i\, x_j}{2} \qquad (i, j = 1, \cdots, n).$$

It follows from the definition of α_{ij} and from (2.1) that the origin is the only

critical point of $\bar{\varphi}$ on or within this sphere, and that it is non-degenerate. Further, its index as a critical point of $\bar{\varphi}$ is the same as its index as a critical point of f.

Finally we shall show that for e sufficiently small and for r between e and $2e$, $\bar{\varphi}$ has no critical points. It will be sufficient to prove that the sum $\bar{\varphi}_{x_i}\bar{\varphi}_{x_i}$ does not vanish for such values of r and e.

By Taylor's formula there exists a neighborhood of the origin, say $r < r_0$, in which

$$f(x) \equiv \frac{\alpha_{ij}\, x_i\, x_j}{2} + \eta(x) \qquad (i, j = 1, \cdots, n),$$

where $\eta(x)$ is an infinitesimal of order greater than two with respect to r. Choose $2e < r_0$. We then have

$$\bar{\varphi}(x) \equiv f(x) - h(r^2/e^2)\,\eta(x), \qquad r < 2e.$$

Hence for $r < 2e$

$$\bar{\varphi}_{x_i} = f_{x_i} - 2\frac{x_i}{e^2}h'\eta - h\eta_{x_i} = \alpha_{ij}\, x_j - \zeta_i(x, e) \qquad (i, j = 1, \cdots, n),$$

where

$$\zeta_i(x, e) \equiv [1 - h(r^2/e^2)]\,\eta_{x_i}(x) - 2\frac{x_i}{e^2}h'(r^2/e^2)\eta(x).$$

We can write

$$\eta(x) \equiv r\xi(x),$$

where $\xi(x)$ is an infinitesimal of order greater than one with respect to r. Let H be the maximum absolute value of $h'(z)$. Then

$$|\,\zeta_i\,| \leqq |\,\eta_{x_i}\,| + 2\frac{2e}{e^2}H\cdot 2e\,|\,\xi\,| \leqq |\,\eta_{x_i}\,| + 8H\,|\,\xi\,|, \qquad r < 2e < r_0.$$

Thus we can write

$$|\,\zeta_i(x, e)| \leqq r\theta_i(x), \qquad r < 2e < r_0,$$

where

$$\lim_{r\to 0}\theta_i(x) = 0.$$

We set

$$x_i = ry_i, \qquad y_i\, y_i = 1.$$

Then

$$\bar{\varphi}_{x_i}\bar{\varphi}_{x_i} = \alpha_{ij}\, x_j\, \alpha_{ij}\, x_j - 2\alpha_{ij}\, x_j\, \zeta_i + \zeta_i\, \zeta_i$$

and

$$\bar{\varphi}_{x_i}\bar{\varphi}_{x_i} \geqq r^2\,\alpha_{ij}\, y_j\,\alpha_{ij}\, y_j - 2r^2\,|\,\alpha_{ij}\, y_j\,|\,\theta_i + r^2\,\theta_i\,\theta_i. \tag{5.1}$$

The form

$$K = \alpha_{ij}\, y_j\, \alpha_{ij}\, y_j$$

is positive definite since $\alpha_{ij}\, y_j$ vanishes only when $(y) = (0)$. For $y_i\, y_i = 1$ the form K has a positive lower bound, say m. But if r is less than a sufficiently small positive constant, say $r_1 < r_0$,

$$|\, \alpha_{ij}\, y_j \,|\, \theta_i < m/2.$$

Hence for $r < 2e < r_1$ it follows from (5.1) that

$$\bar{\varphi}_{x_i}\, \bar{\varphi}_{x_i} \geqq r^2\, \frac{m}{2}.$$

Consequently for $r < 2e < r_1$, the set of equations $\bar{\varphi}_{x_i} = 0$ $(i = 1, \cdots, n)$ is satisfied only when $r = 0$.

The discussion of the critical points of $\bar{\varphi}$ is complete.

It follows at once from its definition that the function $\bar{\varphi}$ is of class C'' in R. We see then that $\bar{\varphi}$ is a function related to f. It is clear that if e is chosen sufficiently small $\bar{\varphi}$ is less than M at interior points of S and equals M on A. Hence $\bar{\varphi}$ satisfies the boundary conditions α relative to the domain S.

If (x^0) were the only critical point of f the function $\bar{\varphi}$ would satisfy the lemma, except possibly for the last clause. The function $\bar{\varphi}$ has been obtained from f by adding a suitable function neighboring (x^0). By adding similar functions for each of a set of non-overlapping spherical neighborhoods of the remaining critical points a function $\bar{\varphi}$ is obtained which does satisfy the lemma, except possibly for the last clause.

We shall now define the function φ demanded in the lemma. The function φ shall be identically equal to $\bar{\varphi}$ except near a critical point. Near such a point we set

$$\varphi \equiv \bar{\varphi} + ch(r^2/e^2),$$

where h, r, and e have the meanings given above, and c is an arbitrarily small positive constant. Thus

$$\varphi \equiv \bar{\varphi} + c, \qquad r \leqq e,$$

$$\varphi \equiv \bar{\varphi}, \qquad r \geqq 2e.$$

The function φ has critical points which are the same as those of $\bar{\varphi}$ in position and index, except possibly when r is between e and $2e$. But for such values of r, for e sufficiently small and for c equal to zero, the sum $\varphi_{x_i}\, \varphi_{x_i}$ does not vanish. The same must be true for c sufficiently small but positive. The number of critical points of f is finite. Hence it is clear that for each critical point we can choose a constant c as above such that φ assumes distinct critical values at distinct critical points.

The function φ is a function related to f. When the c's are chosen sufficiently small φ satisfies the boundary conditions α relative to the domain S. The proof of the lemma is complete.

Throughout the rest of this part of the paper we shall assume that the function f is a function which is related to the given function and satisfies Lemma 1.

ORDINARY VALUES OF f

6. The trajectories orthogonal to the loci, f = const. Let $(x) = (b)$ be an ordinary point of f. Neighboring this point the trajectories orthogonal to the loci f = const. will be represented by solutions of the differential equations

$$\frac{dx_i}{d\tau} = \frac{f_{x_i}}{f_{x_j} f_{x_j}} \qquad (i, j = 1, \ldots, n). \tag{6.1}$$

In a sufficiently small neighborhood of (b) the denominator of the quotient in the right-hand members of (6.1) is not zero. In such a neighborhood these right-hand members possess continuous partial derivatives with respect to each of the x's. Further these right-hand members do not involve τ at all. It follows from the fundamental existence theorems of ordinary differential equations that there exists a positive constant e so small that for any point $(x) = (a)$ in the e neighborhood of (b) and for any constant τ_0 whatsoever, and range of τ within e of τ_0, there exist functions

$$x_i = h_i(a_1, \cdots, a_n, \tau - \tau_0) \qquad (i = 1, \cdots, n) \tag{6.2}$$

which give for constant values of $(a_1, \cdots, a_n)$ and for a variation of τ, a solution of (6.1) that passes through (a) when $\tau = \tau_0$ and is the only such solution of (6.1). Further, for the above described domain the right-hand members of (6.2) possess continuous partial derivatives of at least the first order with respect to each of their arguments.

Along any trajectory

$$\frac{df}{d\tau} \equiv 1,$$

that is, $\tau \equiv f +$ const. But the arbitrariness of τ_0 in (6.2) makes it possible to choose τ arbitrarily at any one point of the trajectory. Hence we can make $\tau \equiv f$ along the trajectory.

7. Topological equivalence of the complexes $f \leqq a, f \leqq b$. Our first theorem is the following.

THEOREM 1. *Let a and b $(a < b)$ be two ordinary values of f in S such that the interval between a and b contains only ordinary values of f. Then the set of points in S satisfying*

$$f \leqq a \tag{7.1}$$

can be put into one-to-one continuous correspondence with the set of points in S satisfying

$$f \leqq b. \tag{7.2}$$

The number of critical values of f is finite. Hence we can choose an ordinary value α of f which is smaller than a, such that there are no critical values of f between α and a. Let H be a set of orthogonal trajectories in which there is just one trajectory passing through each point of the locus

$$f = \alpha. \tag{7.3}$$

Let these trajectories be represented in the form described in §6 in which the value of the parameter τ at each point equals the value of f at that point.

The constants α and b are ordinary values of f between which there is no critical value of f. Hence each trajectory that starts from a point of (7.3) can be continued through an arc corresponding to a range of values of τ from α to b. On these arcs the parameter values $\tau = a$ and $\tau = b$ lie on the loci $f = a$ and $f = b$, respectively. Moreover, through each point $(x) = (c)$ which satisfies the condition

$$\alpha \leqq f(c) \leqq b \tag{7.4}$$

there will pass just one of these arcs. For through each such point there passes an orthogonal trajectory. On this trajectory τ has the value $f(c)$ at (c). A point tracing this trajectory from the point (c) in the sense of decreasing τ must reach a point at which $\tau = \alpha$. Such a point is on (7.3). Hence the trajectory through it is a member of H. Finally it follows from the fundamental existence theorem cited in §6 that there is not more than one of these trajectories through each point satisfying (7.4).

The correspondence whose existence is affirmed in the theorem can now be set up. Observe that the points satisfying (7.1) consist of the sum of the points satisfying

$$f \leqq \alpha \tag{7.5}$$

and those satisfying

$$\alpha < f \leqq a, \tag{7.6}$$

while the points satisfying (7.2) consist of the points satisfying (7.5) together with the points satisfying

$$\alpha < f \leqq b. \tag{7.7}$$

To set up the required correspondence the points satisfying (7.5) are first made to correspond to themselves. The points satisfying (7.6) are then put into correspondence with those satisfying (7.7) by requiring that any point P satisfying (7.6) that arises on the trajectory T through P from a value of τ that divides the closed interval (α, a) in a certain ratio shall correspond to that point satisfying (7.7) that lies on T and arises from a value of τ that divides the closed interval (α, b) in the same ratio.

That the latter correspondence is one-to-one follows from the fact that the trajectories concerned pass through each of the points concerned with one and

only one trajectory through each point. That this correspondence is continuous follows from the nature of the dependence of the points of the trajectories on τ and on their initial points on (7.3). Further the correspondence of the points satisfying (7.5) with themselves clearly joins up in a continuous manner with the correspondence of the points satisfying (7.6) with those satisfying (7.7). Thus the theorem is proved.

CRITICAL POINTS OF INDEX ZERO

8. **The connectivity numbers of the complex $f \leqq m + e^2$, where m is the absolute minimum of f.** Let e be a positive constant so small that m is the only critical value of f less than or equal to $m + e^2$. Under the boundary conditions α and our assumption regarding the function f made at the end of §5, f will take on its absolute minimum at just one interior point P of S. Suppose P is the origin of the space of the variables (x). Neighboring $(x) = (0)$, $f - m$ equals a non-degenerate quadratic form of index zero. Accordingly a suitable non-singular, homogeneous, linear transformation from the variables (x) to variables (y) will bring $f - m$ to the form

$$f(x) - m \equiv y_1^2 + \cdots + y_n^2.$$

Hence for a sufficiently small positive constant e the points in the neighborhood of P which satisfy

$$f(x) \leqq m + e^2$$

will correspond in a one-to-one continuous manner to the points which satisfy

$$y_1^2 + \cdots + y_n^2 \leqq e^2.$$

These points are the points interior to and on an $(n - 1)$-sphere in n-dimensional space. Such a set of points constitutes, in the terms of analysis situs, an n-cell and its boundary. Noting the definitions of the connectivity numbers we have the following theorem.

THEOREM 2. *If m is the absolute minimum of f in S, then for a sufficiently small positive constant e the set of points in S satisfying*

$$f \leqq m + e^2$$

makes up an n-dimensional complex with the connectivity numbers $R_0 = 1, R_i = 0$, $(i = 1, \cdots, n)$.

9. **A critical point of index zero at which f is not equal to its absolute minimum.** Let us suppose now that P is a critical point of f of index zero, at which f does not assume its absolute minimum. Let $f = c$ at P. Let a and b $(a < b)$ be two ordinary values of f between which c is the only critical value. Exactly as in the proof of Theorem 2 we can conclude that for a sufficiently small positive constant e the points of S neighboring P which satisfy

(9.1) $$f \leqq c + e^2$$

constitute an n-cell a_n and its boundary. And if e is chosen small enough this n-cell and its boundary will be distinct from the remaining points, say C_n, which satisfy (9.1).

We shall choose e so small that $c + e^2 < b$. Then $c + e^2$ is an ordinary value of f, and c is the only critical value of f between a and $c + e^2$. It follows that the points of C_n which satisfy

$$a \leqq f \leqq c + e^2$$

are ordinary points of f. For under our assumption regarding f, P is the only critical point at which $f = c$. Using the orthogonal trajectories as in the proof of Theorem 1 we could show that C_n and the complex $f \leqq a$ are homeomorphic. Consequently the connectivity numbers of C_n and of the complex $f \leqq a$ are identical.

If we add the closure of the n-cell a_n to C_n we obtain the entire set of points satisfying (9.1). The connectivity numbers of the domain (9.1) differ from those of C_n only in that the zero$^{\text{th}}$ connectivity number R_0 of the domain (9.1) is one greater than that of C_n. Finally, since both $c + e^2$ and b are ordinary values of f between which there are no critical values, it follows from Theorem 1 that the complex $f \leqq c + e^2$ is homeomorphic with the complex $f \leqq b$.

We have proved the following theorem.

THEOREM 3. *If a and b $(a < b)$ are any two ordinary values of f between which there is just one critical value of f, say c, taken on at a critical point of index zero the connectivity numbers R_i $(i = 0, 1, \cdots, n)$ of the set of points in S which satisfy*

$$f(x) \leqq b$$

differ from the connectivity numbers R_i' of the set of points in S which satisfy

$$f(x) \leqq a$$

only in that

$$R_0 = R_0' + 1.$$

CRITICAL POINTS OF INDEX 1, 2, $\cdots$, n.

10. **Outline of the method to be followed.** In §§10–13 we shall deal with a critical point of f of index k, where k is one of the integers $1, 2, \cdots, n - 1$. We can suppose that the critical value assumed by f at this point is zero. We shall suppose that e is so small a positive constant that there is no critical value of f other than zero on the closed interval $(-e^2, e^2)$. We shall determine the difference between the connectivity numbers of the complex of points in S satisfying

(10.1) $$f \leqq -e^2$$

and the complex of points in S satisfying

$$f \leqq e^2. \tag{10.2}$$

To that end we first show that the complex (10.1) can be put into one-to-one continuous correspondence with the complex (10.2) provided we exclude a properly chosen neighborhood of the given critical point. This will be accomplished with the aid of the orthogonal trajectories already used in §7. The problem will then be resolved into one of determining the difference between the connectivities of the complex (10.2) and of the complex (10.2) with this neighborhood excluded.

11. **A new choice of coordinates.** We suppose the critical point under consideration is the origin. Let a non-singular, linear, homogeneous transformation from our variables (x) to variables (y) be made of such a character that in a neighborhood N of the origin

$$f \equiv -y_1^2 - \cdots - y_k^2 + y_{k+1}^2 + \cdots + y_n^2. \tag{11.1}$$

In the space (y) the trajectories orthogonal to the loci $f = \text{const.}$ will be defined by the differential equations

$$\frac{dy_i}{dt} = \frac{\partial f}{\partial y_i} \qquad (i = 1, \cdots, n). \tag{11.2}$$

On N the equations (11.2) have the form

$$\begin{aligned} \frac{dy_i}{dt} &= -2y_i && (i = 1, \cdots, k) \\ \frac{dy_j}{dt} &= 2y_j && (j = k+1, \cdots, n). \end{aligned} \tag{11.3}$$

We set

$$\begin{aligned} p^2 &= y_1^2 + \cdots + y_k^2, \\ q^2 &= y_{k+1}^2 + \cdots + y_n^2, \end{aligned} \qquad \text{(on } N\text{)} \tag{11.4}$$

where p^2 and q^2 are defined only on N. Let a and e $(a < e)$ be two arbitrarily small positive constants. We shall be concerned with sets of points A_n and E_n in S which satisfy the following conditions:

$$\begin{aligned} &A_n\colon f \leqq -e^2 \\ &E_n\colon -e^2 < f \leqq e^2, \quad p^2 \geqq a^2. \end{aligned}$$

Where p^2 is not defined the inequality $p^2 > a^2$ will be considered as satisfied vacuously. The point-set $A_n + E_n$ is the complex $f \leqq e^2$ with a suitable set of points in the neighborhood of the origin excluded.

12. **The topological equivalence of the complexes A_n and $A_n + E_n$.** We shall prove the following lemma.

LEMMA 2. *For a sufficiently small positive constant e the complexes A_n and $A_n + E_n$ are homeomorphic.*

One restriction on e has already been made in §10. It will be convenient in the proof to restrict e still further. To prove Lemma 2, Lemmas A, B, C, D and E will first be established. Observe that the boundary of E_n is the sum of the three following point-sets:

$$B_1\colon f = e^2, \qquad p^2 > a^2,$$

$$B_2\colon f \leqq e^2, \qquad p^2 = a^2,$$

$$B_3\colon f = -e^2.$$

We suppose e so small, with $a < e$, that B_2 defines a closed point-set interior to an open neighborhood in which (11.1) holds.

LEMMA A. *In the space (y) an orthogonal trajectory which enters E_n by crossing B_2 intersects the boundary of E_n again for the first time on B_3.*

Let $(y) = (b)$ be an arbitrary point of B_2. Then (b) is an ordinary point of f. Hence there is a unique orthogonal trajectory T through (b). It will appear in the course of the proof that T actually crosses B_2. If we choose e sufficiently small the differential equations of T will admit the representation (11.3), at least neighboring (b). Hence near (b) the finite equations of T are

$$y_i = b_i e^{-2t} \qquad (i = 1, \cdots, k),$$

$$y_j = b_j e^{2t} \qquad (j = k+1, \cdots, n),$$

where we have taken $t = 0$ at (b).

As t decreases from zero T enters E_n. For on T near (b)

$$p^2 = \sum_{i=1}^{k} b_i^2 e^{-4t}.$$

Hence at (b)

$$\frac{dp^2}{dt} = -4 \sum_{i=1}^{k} b_i^2 = -4a^2 < 0.$$

That is, as t decreases from zero p^2 increases from a^2. Further, on T near (b)

$$\frac{df}{dt} = \frac{\partial f}{\partial y_i}\frac{dy_i}{dt} = 4y_i y_i > 0 \qquad (i = 1, \cdots, n). \tag{12.1}$$

Hence as t decreases from zero f decreases. But at (b), $f \leqq e^2$, and we see that as t decreases from zero T actually enters E_n.

Suppose now that T has the representation in which the value of the parameter τ at any point of T is equal to the value of f at that point. It follows from (12.1) that τ is decreasing with t at (b). Let T be traced in the sense of decreasing τ. We note the following.

(i) The trajectory T does not return to B_2. For suppose T intersected B_2 again, say at Q. At Q the parameter t would be well defined. Since τ is decreasing t would be decreasing at Q. We would then have an orthogonal trajectory leaving E_n with decreasing t. This would be a contradiction of the result just obtained.

(ii) The trajectory T cannot intersect B_1. For on B_1 the parameter value τ would be e^2. But at (b), $\tau \leqq e^2$, and τ decreases along T. Hence statement (ii) is true.

(iii) As τ decreases T must eventually leave E_n. The value of e has been chosen so small that zero is the only critical value of f on the closed interval $(-e^2, e^2)$. Moreover the only critical point at which $f = 0$ is the origin. But the origin is not a point of E_n. Hence we can continue T in the sense of decreasing τ until $\tau < -e^2$. It follows from the definition of E_n that T will then have left E_n.

The combination of (i), (ii) and (iii) proves the lemma.

A similar lemma is the following.

LEMMA B. *Every orthogonal trajectory which enters E_n by crossing B_1 intersects the boundary of E_n again for the first time on B_3.*

Let Q be an arbitrary point of B_1. Then Q is an ordinary point of f. Through Q there will be a unique trajectory T. At Q, $f = e^2$, while at interior points of E_n, $f < e^2$. Hence as τ decreases T enters E_n at Q. We see as in the proof of Lemma A that T cannot return to B_1 nor intersect B_2. But T eventually leaves E_n and hence intersects B_3 as stated.

LEMMA C. *Every orthogonal trajectory which enters E_n by crossing B_3 intersects the boundary of E_n again for the first time on B_1 or B_2.*

The proof of this lemma is similar to that of Lemma A.

LEMMA D. *There exists a one-to-one continuous correspondence between the points P on B_3 and the points Q on $B_1 + B_2$ in which corresponding points lie on the same orthogonal trajectory T.*

Lemmas A, B and C establish the truth of this lemma except for the continuity of the correspondence. Let Q_0 be any point on $B_1 + B_2$. The trajectory T through Q_0 intersects B_3 in P_0, the correspondent of Q_0 on B_3. In the neighborhood of Q_0 there is no critical point of f. Hence if Q is any point in this neighborhood it follows from the fundamental existence theorems for differential equations that a point on the trajectory through Q varies continuously with Q and the value of a parameter $\tau = f$ on this trajectory. In particular Q may vary continuously on $B_1 + B_2$ and τ may be taken as $-e^2$. The resulting point P on T will vary continuously on B_3.

Conversely Q will vary continuously on $B_1 + B_2$ as P varies continuously on B_3. This fact is a consequence of the following theorem. If there exists a one-to-one correspondence between a closed set of points Q and a set of points P in which P varies continuously with Q then Q varies continuously with P. See v. Kerékjártó (7) p. 34. The lemma is proved.

We conclude with the following lemma.

LEMMA E. *The arcs of the trajectories T of Lemma D which join points P and Q on B_3 and $B_1 + B_2$, respectively, include one and only one arc through each point of E_n.*

Let M be any interior point of E_n. Then through M there is one and only one orthogonal trajectory T'. The parameter value τ on T' at M lies between $-e^2$ and e^2. A point tracing T' from M in the sense of decreasing τ must reach a point P_1 on B_3, and in the sense of increasing τ a point Q_1 on $B_1 + B_2$. It follows from Lemma D that P_1 and Q_1 correspond in the sense of that lemma. The lemma is proved.

Proof of Lemma 2. We can now set up the homeomorphism of Lemma 2. Let α be an ordinary value of f which is smaller than $-e^2$ and such that no critical value of f lies between α and $-e^2$. Observe that the complex A_n consists of the points which satisfy

$$f \leqq \alpha \tag{12.2}$$

and the set G_n of points which satisfy

$$\alpha < f \leqq -e^2.$$

To set up the required correspondence the domain $f \leqq \alpha$ is first made to correspond with itself. The domain G_n can be put into correspondence with the domain $G_n + E_n$ in the following manner. Let P and Q be points on B_3 and $B_1 + B_2$, respectively, which lie on the same trajectory T. Let T be continued in the sense of decreasing τ. Let R be the point on T at which $\tau = \alpha$. Let the arc of the trajectory T between R and Q be denoted by T'. It is readily seen that Lemmas D and E hold if the points R and Q and domain $G_n + E_n$ replace the points P and Q and the domain E_n of these lemmas, respectively. The value of τ on T' at P is $-e^2$. Let the value of τ on T' at Q be denoted by τ'. The required correspondence shall be one in which a point on T' at which τ divides the closed interval $(\alpha, -e^2)$ in a certain ratio corresponds to that point on T' at which τ divides the closed interval (α, τ') in the same ratio.

The correspondence between G_n and $G_n + E_n$ joins on continuously with the correspondence of $f \leqq \alpha$ with itself. It follows as in the proofs of Lemmas E and D that our correspondence is one-to-one and continuous.

The proof of Lemma 2 is complete.

13. **Incidence relations between D_n and the complex $f \leqq e^2$ with the new points of D_n excluded.** Let D_n denote the set of points which satisfy

$$f \leqq e^2, \qquad p^2 \leqq a^2,$$

where e, p, and a have the same meanings as in the last section. We shall show that the points of D_n constitute an n-cell and its boundary. We shall also determine the incidence relations between D_n and the closure of the complex $f \leqq e^2$ with D_n excluded.

Let $(c) = (c_1, \cdots, c_n)$ be the coordinates of a variable point on the unit $(n-1)$-sphere S_{n-1} with center at the point $(y) = (0)$. By the methods of elementary analytical geometry it can readily be shown that the ray γ joining the origin to the point (c) intersects the boundary of D_n in a unique point (y) which varies continuously with (c). Moreover, the distance from the origin to the point (y), namely $r = \sqrt{y_i y_i}$, is a continuous function of (c).

The points interior to and on S_{n-1} constitute an n-cell and its boundary. We shall show that the points of D_n constitute an n-cell and its boundary by setting up a homeomorphism between the points of D_n and the points interior to and on S_{n-1}. Let an arbitrary ray γ from the origin intersect the boundary of D_n in the point $(y)_\gamma$ and S_{n-1} in the point $(c)_\gamma$. Let γ' and γ'' be the segments of γ which lead from the origin to the points $(y)_\gamma$ and $(c)_\gamma$ respectively.

The correspondence shall be one in which points on γ' and γ'' correspond which divide γ' and γ'' respectively in the same ratio. It follows from the remarks made above that the correspondence is a homeomorphism.

It is convenient to think of each point as having two sets of identical coordinates (y) and (z). We set

$$u^2 = z_1^2 + \cdots + z_k^2,$$
$$v^2 = z_{k+1}^2 + \cdots + z_n^2.$$

The following lemma can be established by methods similar to those indicated above.

LEMMA 3. *There exists a homeomorphism between the points (y) of D_n and the points (z) satisfying*

$$u^2 \leqq a^2, \qquad v^2 \leqq b^2, \tag{13.1}$$

where $b^2 = a^2 + e^2$, in which points on the boundary of D_n which satisfy

$$f \leqq e^2, \qquad p^2 = a^2 \tag{13.2}$$

correspond to points on the boundary of the domain (13.1) *which satisfy*

$$v^2 \leqq b^2, \qquad u^2 = a^2$$

while points on the boundary of D_n which satisfy

$$f = e^2, \qquad p^2 < a^2 \tag{13.3}$$

correspond to points on the boundary of the domain (13.1) *which satisfy*

$$v^2 = b^2, \qquad u^2 < a^2.$$

We observe that the points (y) of D_n satisfy

$$f \leqq e^2 \tag{13.4}$$

and are separated from the remaining points (y) which satisfy (13.4) by those boundary points of D_n which satisfy (13.2). These boundary points of D_n will

be termed *the old boundary points of* D_n, while the remaining boundary points of D_n, namely those which satisfy (13.3), will be termed *the new boundary points of* D_n. The set of all points of D_n other than the old boundary points will be called *the new points of* D_n.

The points (z) which satisfy (13.1) may be represented by a pair of points (P, Q) of which P is a point in a k-dimensional space with coordinates $(z_1, \cdots, z_k)$ satisfying $u^2 \leqq a^2$, and Q is a point in an $(n - k)$-dimensional space with coordinates $(z_{k+1}, \cdots, z_n)$ satisfying $v^2 \leqq b^2$. Let the $(k - 1)$-sphere $u^2 = a^2$ in the space of the points P be denoted by S_{k-1}, and the $(n - k - 1)$-sphere $v^2 = b^2$ in the space of the points Q be denoted by S_{n-k-1}. With the above conventions the results of Lemma 3 may be translated into the following lemma.

LEMMA 4. *The points of D_n make up an n-cell and its boundary. The set D_n can be put into one-to-one continuous correspondence with the set of pairs of points (P, Q) obtained by combining an arbitrary point P interior to or on a $(k - 1)$-sphere S_{k-1} with an arbitrary point Q interior to or on an $(n - k - 1)$-sphere S_{n-k-1}. In this correspondence points on the old boundary of D_n correspond to those pairs (P, Q) that are obtained by combining an arbitrary point P on S_{k-1} with an arbitrary point Q interior to or on S_{n-k-1}, while the points on the new boundary of D_n correspond to those pairs (P, Q) that are obtained by combining an arbitrary point P interior to S_{k-1} with an arbitrary point Q on S_{n-k-1}.*

We can now complete Lemma 2 as follows.

LEMMA 5. *If e is a sufficiently small positive constant the complex $f \leqq -e^2$ can be put into one-to-one continuous correspondence with the complex, say C_n', which consists of the points satisfying $f \leqq e^2$ with the new points of D_n excluded. The remaining points of D_n, namely the points on the old boundary of D_n, are on the boundary of C_n'.*

14. **Differences between the connectivity numbers of the complex $f \leqq -e^2$ and the complex $f \leqq e^2$.** At this point the reader is referred to the original paper of which this paper is an extension. See Morse (9). Sections 14, 15, and 16 of that paper can be read without any change except for the numbering of the lemmas and for the observation that in that paper Veblen's definition of the connectivity numbers is used.

We combine the theorems of §§15 and 16 of the earlier paper in the following theorem.

THEOREM 4. *Let c be a critical value assumed by f at a critical point of index k, $0 < k \leqq n$; then for a sufficiently small positive constant e the differences $\Delta R_i = R_i - R_i'$ $(i = 0, 1, \cdots, n)$ between the connectivity numbers R_i of the complex $f \leqq c + e^2$ and the connectivity numbers R_i' of the complex $f \leqq c - e^2$ are all zero, except that either*

Case I. $$\Delta R_k = 1$$

or else

Case II. $\Delta R_{k-1} = -1.$

If $k = n$, Case II *always occurs.*

We shall term a critical point of index k $(k = 0, 1, \cdots, n)$ for which $\Delta R_k = 1$, a critical point of index k of *increasing type*, and one for which $\Delta R_{k-1} = -1$, a critical point of index k of *decreasing type*. For $k = 1, \cdots, n-1$ it can be shown by simple examples that both of these types are possible.

GENERAL RELATIONS BETWEEN ALL THE CRITICAL POINTS

15. In this section we shall derive the relations between the numbers M_i of the critical points of index i of f in S, $(i = 0, 1, \cdots, n)$ and the connectivity numbers R_j of S, $(j = 0, 1, \cdots, n)$.

We first prove the following theorem.

THEOREM 5. *If m_i^+ and m_i^- are the total numbers of critical points of index i of f in S of increasing and decreasing types respectively, and if R_j is the j^{th} connectivity number of S, then under the boundary conditions α*

$$R_i = m_i^+ - m_{i+1}^- \qquad (i = 0, 1, \cdots, n),$$

where $m_{n+1}^- = 0$.

It will be convenient to say that the null complex has null connectivity numbers. Let $c_0 < c_1 < \cdots < c_r$ be the complete set of critical values of f in S. For convenience let c_{r+1} be the absolute maximum M of f in S. It is possible to choose a positive constant e so small that no two of the closed intervals

$$(c_j + e^2, c_{j+1} - e^2) \qquad (j = 0, 1, \cdots, r)$$

contain a common value. None of these closed intervals contains a critical value of f. Hence it follows from Theorem 1 that the connectivity numbers of the complexes $f \leqq c_{j+1} - e^2$ and $f \leqq c_j + e^2$ are identical. The same is true for the complexes $f \leqq c_{r+1}$ and $f \leqq c_r + e^2$. Further, on each of the closed intervals

$$(c_j - e^2, c_j + e^2) \qquad (j = 0, 1, \cdots, r)$$

c_j is the only critical value of f. Hence as c takes on the successive values

$$(15.1) \quad c_0 - e^2, \quad c_0 + e^2, \quad c_1 - e^2, \quad c_1 + e^2, \cdots, \quad c_r - e^2, \quad c_r + e^2, \quad c_{r+1}$$

the only changes in the connectivity numbers of the complexes $f \leqq c$ occur as c passes from $c_j - e^2$ to $c_j + e^2$, $(j = 0, 1, \cdots, r)$.

The differences between the connectivity numbers of the complex $f \leqq c_j + e^2$ and the connectivity numbers of the complex $f \leqq c_j - e^2$ $(j = 0, 1, \cdots, r)$ are given in Theorems 2, 3 and 4. It is now immediately verifiable that the sums of these changes as c passes through the values (15.1) from $c_0 - e^2$ to c_{r+1} are given by the equations of the theorem.

The following theorem gives the answer to one of the fundamental questions with which this paper is concerned.

THEOREM 6. *If M_i $(i = 0, 1, \cdots, n)$ is the number of critical points of index i of f in S, and if R_j $(j = 0, 1, \cdots, n)$ is the j^{th} connectivity number of S, then under the boundary conditions α the following relations hold:*

$$R_0 \leqq M_0$$
$$R_0 - R_1 \geqq M_0 - M_1$$
$$R_0 - R_1 + R_2 \leqq M_0 - M_1 + M_2$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$R_0 - R_1 + R_2 - \cdots + (-1)^n R_n = M_0 - M_1 + M_2 - \cdots + (-1)^n M_n.$$

It follows from the definition of m_i^+ and m_i^- that

$$M_i = m_i^+ + m_i^- \qquad (i = 0, 1, \cdots, n).$$

The numbers m_i^+ and m_i^- are all positive integers or zero. In particular $m_0^- = m_n^+ = 0$. The proof of Theorem 6 consists merely of verifying that the relations of the theorem follow from the equations of Theorem 5.

COROLLARY. *Under the hypotheses of the theorem*

$$M_i \geqq R_i \qquad (i = 0, 1, \cdots, n).$$

There are then at least

$$R_0 + R_1 + \cdots + R_n$$

critical points of f in S.

The preceding theory can equally well be developed for the case of functions defined on closed regular manifolds locally representable in terms of n parameters (x). The degeneracy of a critical point or its index will be invariant of non-singular changes of parameters. In Theorem 4, Case II can now occur when $k = n$. There are, of course, no boundary conditions. The proof of Theorem 6 goes through essentially as before. See Morse (12) Chapter VI.

THE GENERAL CASE $f_N > 0$

16. **The boundary conditions β.** We shall now replace the boundary conditions α by a set of boundary conditions which we shall term boundary conditions β.

Let Σ be any open n-dimensional region whose closure $\bar{\Sigma}$ lies on the open region R of §2. The boundary B of Σ consists of a bounded closed set of points lying in the space of the variables $(x_1, \cdots, x_n)$. The *boundary conditions β* are the following:

Iβ. *The points on B in the neighborhood of any particular point (x^0) on B shall satisfy a relation of the form*

$$\Phi(x_1, \cdots, x_n) = 0,$$

where $\Phi(x)$ is a real, single-valued function of $(x_1, \cdots, x_n)$ of class C''' neighboring (x^0), at least one of whose first partial derivatives does not vanish at (x^0).

IIβ. *The domain* Σ *shall contain all the critical points of* f.

IIIβ. *At each point of* B *the directional derivative* f_N *of* $f(x)$ *along the normal to* B *in the sense that leads from points on* Σ *to points not on* Σ *shall be positive.*

We shall refer to these conditions more explicitly as the boundary conditions β on f *relative* to Σ.

17. **New coordinates neighboring** B. It will be convenient to introduce new coordinates for the points of R neighboring B. We make the following definitions. If a transformation from the coordinates (x) to coordinates (ξ) is defined neighboring a point (x^0) by functions $\xi_i(x)$ which are of class C'' and if the jacobian of the transformation does not vanish at (x^0) the coordinates (ξ) will be termed *admissible* neighboring (x^0). An $(n-1)$-parameter family of straight lines will be said to form a *field* neighboring a point (x^0) if there is one and only one straight line of the family through each point of a sufficiently small neighborhood of (x^0). We shall prove the following lemma.

LEMMA 6. *Let* (x^0) *be an arbitrary point on* B. *The normals to* B *neighboring* (x^0) *form a field neighboring* (x^0).

It follows from Iβ that one of the derivatives Φ_{x_i} is not zero at (x^0). Suppose this derivative is Φ_{x_n}. Then in a neighborhood of (x^0) the equation $\Phi(x) = 0$ can be solved for x_n in the form

$$x_n = g(x_1, \cdots, x_{n-1}),$$

where g is a function of class C'''. Upon setting

$$x_j = u_j, \qquad x_j^0 = u_j^0 \qquad (j = 1, \cdots, n-1)$$

we see that B can be represented neighboring (x^0) in the form

$$x_i = g_i(u_1, \cdots, u_{n-1}) \qquad (i = 1, \cdots, n), \tag{17.1}$$

where the matrix of the first partial derivatives of the functions g_i has the rank $n-1$ at the point (u^0). Since B is a bounded closed set of points it follows from the Heine-Borel theorem that B can be covered with a finite set of neighborhoods in each of which B has a representation of the form (17.1).

Let $b_k(u)$ be the determinant obtained from the matrix

$$\left\| \frac{\partial g_i}{\partial u_j} \right\| \qquad \begin{matrix} (i = 1, \cdots, n) \\ (j = 1, \cdots, n-1) \end{matrix} \tag{17.2}$$

by deleting the k^{th} row. Then the set $(-1)^k b_k(u)$ $(k = 1, \cdots, n)$ is a set of direction numbers of the normal to B at the point (u). The direction cosines of the exterior normal are

$$a_i(u) = \frac{(-1)^i e\, b_i(u)}{\sqrt{b_j(u)\, b_j(u)}} \qquad (i, j = 1, \cdots, n),$$

where e is taken as 1 or -1 according as the set $(-1)^i b_i(u)$ gives direction numbers of the exterior or interior normal. The equations of the normal to B at (u) are

$$x_i = x_i(u, s) = g_i(u) + sa_i(u) \qquad (i = 1, \cdots, n), \tag{17.3}$$

where s is the arc-length along the normal measured from B. For points on Σ we see that $s < 0$.

To prove the lemma it will be sufficient to show that the jacobian of the right-hand members of (17.3) with respect to the variables (u, s) is not zero when $(u, s) = (u^0, 0)$. This jacobian is the determinant

$$J = \begin{vmatrix} \dfrac{\partial g_1}{\partial u_1} & \cdots & \dfrac{\partial g_1}{\partial u_{n-1}} & a_1 \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \dfrac{\partial g_n}{\partial u_1} & \cdots & \dfrac{\partial g_n}{\partial u_{n-1}} & a_n \end{vmatrix}_{(u) = (u^0)}. \tag{17.4}$$

The matrix of the first $n - 1$ columns of J has the rank $n - 1$. Hence the vanishing of J would imply the dependence of the last column upon the other columns. But the last column gives the direction cosines of the normal to B at (u^0), while the other columns give direction numbers of $n - 1$ linearly independent directions tangent to B at (u^0). The normal direction is not dependent upon the tangent directions. Hence J does not vanish. The proof of Lemma 6 is complete.

The equations (17.3) define an admissible change of coordinates from the coordinates (x) neighboring (x^0) to coordinates (u, s) neighboring $(u^0, 0)$. We shall further transform the coordinates (u, s) neighboring $(u^0, 0)$ into coordinates (u, v) such that

$$\begin{aligned} u_i &= u_i \qquad (i = 1, \cdots, n - 1), \\ v &= f[x(u, s)] - f[(x(u, 0)]. \end{aligned} \tag{17.5}$$

This transformation is of class C''. One verifies the fact that for $s = 0$

$$\frac{\partial v}{\partial s} = f_N > 0.$$

From this fact it follows that the jacobian of this transformation is not zero at the point (u^0) on B. The parameters (u, v) thus form an admissible set of coordinates neighboring the point (x^0).

Corresponding to the point (u^0) on B there exists a positive constant η and on B a neighborhood N of (u^0) such that the normals to B at the points of N have the following properties. The segments of these normals on which $|v| \leqq \eta$ cover a neighborhood N^* of (x^0) in a one-to-one manner while at each point of N^* the equations (17.3) and (17.5) define a transformation with non-vanishing jacobian. Sets (u, v) for which (u) lies on N and $|v| \leqq \eta$ may thus be regarded as admissible coordinates for the neighborhood N^*. It follows from

the Heine-Borel theorem that B can be covered with a finite number of neighborhoods

$$N_1, \cdots, N_\mu \tag{17.6}$$

such as N, and that accordingly one choice of the constants η can be made for all of these neighborhoods. It follows from the definition of η that f has no critical points on any of the neighborhoods N^* corresponding to the neighborhoods (17.6).

18. **Redefinition of the function f.** Let e_1, e_2, e_3 be three constants such that $-\eta < e_1 < e_2 < e_3 < 0$. The points on the normals to B at which v has the values e_1, e_2, e_3 constitute three loci T_1, T_2, T_3, respectively. Let N be any one of the neighborhoods (17.6) and let the corresponding set of parameters on B be (u). Let $F(u)$ be the value of f at the point (u) on B.

At the points between T_1 and B we define a function $\psi(u_1, \cdots, u_{n-1}, v)$ as follows. Let G be a constant greater than the absolute maximum of f on $\bar{\Sigma}$. Let $H(u_1, \cdots, u_{n-1})$ be the function defined by the equation

$$F(u) + e_3 + H(u)\,[e_3 - e_2]^4 = G. \tag{18.1}$$

The function $H(u)$ is of class C'' on N. We now set

$$\psi(u, v) \equiv F(u) + v, \qquad e_1 < v \leqq e_2,$$

$$\psi(u, v) \equiv F(u) + v + H(u)\,[v - e_2],^4 \qquad e_2 < v < 0.$$

The function ψ is of class C''. On T_3, $v = e_3$ and by virtue of (18.1), $\psi \equiv G$. We observe that $H(u)$ is positive. Hence ψ_v is positive. Thus ψ has no critical points.

We now redefine the function f on Σ. Let the new function be $L(x_1, \cdots, x_n)$. The function $L(x)$ shall be identically equal to $f(x)$ on Σ except at the points (x) at which $e_1 < v < 0$. At each of the latter points $L(x)$ shall be equal to the value of ψ at the corresponding point (u, v). It follows from the properties of ψ that $L(x)$ is of class C'' on Σ.

19. **The relation between the critical points of f and the connectivities of $\bar{\Sigma}$.** The function $L(x)$ defined in the preceding section is a function related to f in the sense of §5, provided we restrict f to its values on Σ. Let S denote the closed domain of points (x) which satisfy the condition

$$L(x) \leqq G.$$

The domain S is bounded by T_3.

We state the following lemma.

LEMMA 7. *The domain S is homeomorphic with $\bar{\Sigma}$.*

This lemma is readily proved by using the normals to B in the way in which the orthogonal trajectories were used in the proof of Theorem 1.

We summarize the preceding results in the following theorem.

THEOREM 7. *Suppose $f(x)$ satisfies the boundary conditions β relative to the region Σ. Then there exists a function $L(x)$ related to $f(x)$ on Σ and a sub-domain S of Σ homeomorphic with $\bar{\Sigma}$ such that $L(x)$ satisfies the boundary conditions α relative to S.*

Since S is homeomorphic with $\bar{\Sigma}$ the connectivity numbers of S and $\bar{\Sigma}$ are the same. Hence Theorems 6 and 7 combine to give the following theorem.

THEOREM 8. *Under the boundary conditions β the relations between the numbers M_i of critical points of index i of f on Σ and the connectivity numbers R_j of $\bar{\Sigma}$ $(i, j = 0, 1, \cdots, n)$ are the same as those in Theorem 6.*

We state the following theorem.[2] No proof will be given in this place.

THEOREM 9. *Let there be given a set of integers M_i and R_j, positive or zero, satisfying the relations of Theorem 6 with $R_0 \geqq 1$. In euclidean n-space there exists a bounded, open region R and on R a non-degenerate function f of class C'', together with a constant M, such that the domain of points $f \leqq M$ is closed and interior to R, contains exactly M_i critical points of f of index i, and possesses a j^{th} connectivity number R_j.*

Accordingly, under the type of boundary conditions we have admitted there are in general no relations other than those of Theorem 6 between the integers M_i and R_j.

GENERAL BOUNDARY CONDITIONS

20. **The boundary conditions γ.** We come to the general boundary conditions. Let the region Σ and its boundary B be defined as in §16.

Under boundary conditions γ we retain conditions Iβ *and* IIβ *of boundary conditions β.*

Let P be any point on B. It follows from boundary condition Iβ that B can be represented neighboring P in the form

$$x_i = g_i(u_1, \cdots, u_{n-1}) \qquad (i = 1, \cdots, n),$$

where the functions g_i are of class C''' neighboring the point (u^0) that determines P, and where the matrix of the first partial derivatives of the functions g_i is of rank $n - 1$ at (u^0). Parameters (u) in such a local representation of B will be termed *admissible*. Let φ be the function defined by f on B. We shall represent φ in terms of local parameters of B, using however only those parameters of B which we have termed admissible. The condition that φ have a critical point and that that critical point be non-degenerate of index k will be independent of the particular set of admissible parameters used locally to represent B.

---

[2] The theory of critical points and its applications was studied by the seminar in the University of Göttingen in 1932–33. Dr. John of this seminar independently obtained results similar to those of Theorem 9. See John (6).

The reader who compares the definition of the region D_n of §13 with the definition given by Morse in the earlier paper will notice that in the present paper $a < e$, and in the earlier paper $e < a$. This simplification was suggested by Herr Wittmeyer of this same seminar.

Under boundary conditions γ the third condition is as follows.

IIIγ. *The boundary function φ defined by f on B shall be non-degenerate.*

Let f_N be the directional derivative of f along the normal to B in the sense that leads from points within Σ to points not within Σ. The set of points on B at which $f_N < 0$ will be termed the *negative boundary* of Σ.

21. **A modification of f neighboring B.** Let (u) be an admissible coordinate system on B covering a neighborhood N of a point (u^0) on B. Let (x^0) be the point on B determined by (u^0). Neighboring the point (x^0) we make a transformation of the form (17.3) from the coordinates (x) to coordinates (u, s). The new coordinates (u, s) will form an admissible coordinate system for a neighborhood N^* of (x^0). A particular choice of N^* is the domain (u, s) for which (u) is on N and $|s|$ is at most a sufficiently small positive constant s_1. We observe that s_1 can be chosen independently of the point (x^0) on B and the particular admissible set of coordinates (u) used to cover the neighborhood N on B. In particular s_1 can be chosen so small that f has no critical points on any of the corresponding neighborhoods N^*, and such that the points on the normals to B for which $0 < s \leqq s_1$ are all points of R. We term such a choice of s_1 *admissible*.

Let $s_0 < s_1$ be an arbitrarily small positive constant. Let Σ_0 denote the points of Σ together with the points neighboring B for which $0 \leqq s < s_0$. Let Σ_1 denote the points of Σ together with the points neighboring B for which $0 \leqq s < s_1$. Let B_0 be the boundary of Σ_0.

Let $F(u, s)$ be the value of $f(x)$ at the point (u, s) determined by the point (x) neighboring (x^0). Let M be a positive constant greater than the absolute value of the directional derivative of f on the normals to B at points on these normals for which $0 \leqq s \leqq s_1$. At the points (u, s) for which $0 \leqq s < s_1$ we set

$$\psi(u, s) \equiv F(u, s) + \frac{s^3}{3} \cdot \frac{M}{s_0^2}. \tag{21.1}$$

The function ψ is of class C''.

The normals to B are also normals to B_0. On B_0, $s = s_0$. Hence from (21.1) we see that

$$\psi_s = f_N + M > 0 \tag{21.2}$$

on B_0.

On Σ_1 we replace the function $f(x)$ by a function $\Psi(x)$ defined as follows. On Σ the function Ψ shall be identically equal to f. On each neighborhood N^* of a point (x^0) on B representable by admissible coordinates (u, s), at points for which $0 \leqq s < s_1$, we set

$$\Psi(x) = \psi(u, s) \tag{21.3}$$

where $\psi(u, s)$ has just been defined. It follows from (21.2) that the normal derivative Ψ_N of Ψ on B_0 is positive without exception.

22. **The critical points of Ψ.** It follows from the definition of the function ψ that on $\bar{\Sigma}$ the function Ψ has the same critical points as f with the same indices. The nature of the critical points of Ψ at the remaining points of Σ_1 will be described in Lemma 9 below.

We shall first establish the following lemma.

LEMMA 8. *At a critical point P of the boundary function φ, $f_N \neq 0$.*

Let a neighborhood of P be referred to admissible coordinates (u, s) as in the preceding section. Since P is a critical point of φ we know that

$$\varphi_{u_i} = F_{u_i} = 0 \qquad (i = 1, \cdots, n-1) \tag{22.1}$$

at P. It follows from the nature of the coordinate s that

$$F_s = f_N$$

at P. Hence the vanishing of f_N at P would imply that P were a critical point of f contrary to boundary condition IIγ. The lemma follows directly.

Let

$$P_1, \cdots, P_\mu \tag{22.2}$$

be the critical points of φ on the "negative boundary" of Σ. With the point P_i we associate a neighborhood N_i of P_i on B admissibly represented by parameters (u). We suppose the neighborhoods N_i are so small that their closures are distinct. In the preceding section we have termed certain values of s_1 admissible. Let s_1 be an arbitrary one of these values and let N_i^* be the domain of points (u, s) for which (u) is on N_i and $0 \leqq s < s_1$. The domains N_i^* are distinct.

We continue with the following lemma.

LEMMA 9. *If the preceding constant s_1 and the positive constant $s_0 < s_1$ are chosen sufficiently small the critical points of the function Ψ on the subdomain of Σ_1 for which $0 < s < s_1$ form a set*

$$K_1, \cdots, K_\mu$$

of non-degenerate critical points of which K_i lies on N_i^* and has the same index as P_i.

In terms of the coordinates (u, s) representing the domain N_i^* a critical point of Ψ on N_i^* must satisfy the equations

$$\psi_{u_h}(u, s) = F_{u_h}(u, s) = 0 \qquad (h = 1, \cdots, n-1). \tag{22.3}$$

By virtue of (22.1) the equations (22.3) have an initial solution at P_i. Let the jacobian of the functions ψ_{u_h} be denoted by $A_{n-1}(u, s)$. We have

$$A_{n-1}(u, s) = |\psi_{u_h u_j}| = |F_{u_h u_j}| \qquad (h, j = 1, \cdots, n-1). \tag{22.4}$$

The critical points of φ are non-degenerate. Hence $A_{n-1}(u, s)$ does not vanish at P_i. It follows from the implicit function theorems that for s_1 sufficiently small the solutions of equations (22.3) in N_i^* consist of the points on a curve g_i of the form

$$u_h = u_h^{(i)}(s) \qquad (h = 1, \cdots, n-1),$$

where the functions $u_h^{(i)}(s)$ are single-valued and of class C' for $0 \leqq s < s_1$. The curve g_i issues from the point P_i. It is independent of the constant s_0 and of the constant s_1 except for the fact that s_1 determines the interval on which the functions $u_h^{(i)}(s)$ are defined.

The only critical points of Ψ on N_i^* are on g_i. Let Q_i be the point in which g_i meets B_0, that is, the point on g_i at which $s = s_0$. We shall show that for s_0 sufficiently small Ψ has just one critical point on g_i and that this point lies between P_i and Q_i. To this end we set

$$\theta_i(s) \equiv \psi_s[u^{(i)}(s), s].$$

The function Ψ has a critical point on g_i when and only when $\theta_i(s) = 0$.

We continue the proof with the following statement.

If s_0 is sufficiently small the function $\theta_i(s)$ vanishes once and only once on the interval $0 < s < s_0$ and does not vanish on the interval $s_0 \leqq s < s_1$.

Observe that

$$\theta_i(s) = F_s[u^{(i)}(s), s] + Ms^2/s_0^2. \tag{22.5}$$

Recall that s_0 is a positive constant less than s_1. Let r be a positive constant less unity. For s on the interval

$$0 \leqq s \leqq rs_0 \tag{22.6}$$

we see that

$$\theta_i(s) \leqq F_s[u^{(i)}(s), s] + r^2M.$$

Accordingly for s_0 on the interval $0 < s < s_1$ there exists a choice of r independent of s_0 so small that

$$\theta_i(s) < 0, \qquad 0 \leqq s \leqq rs_0. \tag{22.7}$$

Recall that $\theta_i(s_0) > 0$. We see that $\theta_i(s)$ vanishes at least once on the interval

$$rs_0 < s < s_0.$$

We note that $\theta_i'(s)$ depends upon s_0 only through the term $2Ms/s_0^2$. Accordingly for a further suitable diminution of s_0 (r fixed)

$$\theta_i'(s) > 0, \qquad rs_0 \leqq s < s_1. \tag{22.8}$$

It follows from (22.7) and (22.8) that $\theta_i(s)$ vanishes at most once on the interval $0 < s < s_0$. Hence $\theta_i(s)$ vanishes just once on the interval $0 < s < s_0$. From the fact that $\theta_i(s_0) > 0$ and from (22.8) it follows that $\theta_i(s)$ does not vanish on the interval $s_0 \leqq s < s_1$.

We see then that Ψ has just one critical point K_i on N_i^*. To prove that K_i is a non-degenerate critical point it will be sufficient to show that the hessian of ψ does not vanish at K_i. This hessian is the bordered determinant:

$$(22.9) \qquad H(u, s) = \begin{vmatrix} A_{n-1}(u, s) & \psi_{uhs}(u, s) \\ \psi_{suj}(u, s) & \psi_{ss}(u, s) \end{vmatrix} \qquad (h, j = 1, \cdots, n-1),$$

where $A_{n-1}(u, s)$ is the jacobian (22.4). The only element of $H(u, s)$ which depends upon s_0 is ψ_{ss}. The critical point K_i is determined by a value of s between rs_0 and s_0 where r is a positive constant independent of s_0. It thus follows from (21.1) that s_0 can be chosen so small that ψ_{ss} will be positive at K_i, and in fact will be so large that the term $A_{n-1}\psi_{ss}$ will dominate the value of $H(u, s)$ at K_i, making $H(u, s)$ different from zero there. Hence K_i will be a non-degenerate critical point of Ψ.

To determine the index of K_i we employ the theory of indices of quadratic forms. See Dickson (5) pp. 81–88. Consider the sequence of determinants:

$$(22.10) \qquad A_q = |\psi_{uhuj}| \qquad (h, j = 1, \cdots, q; q = 0, 1, \cdots, n-1)$$

where $A_0 = 1$. Since $A_{n-1} \neq 0$ at P_i it follows from the theory of indices that we can number the variables $u_1, \cdots, u_{n-1}$ in such a way that in the sequence (22.10) no two consecutive terms are zero. If this is done the index k_i of P_i as a critical point of φ will equal the number of changes in sign in the sequence (22.10) evaluated at P_i. None of the terms of (22.10) depends upon s_0. Hence we can choose s_0 so small, and consequently K_i so near P_i, that at K_i the sequence (22.10) will still have exactly k_i changes of sign. We have seen that when s_0 is sufficiently small the signs of A_{n-1} and $H(u, s)$ are the same at K_i. Hence for a proper choice of s_0 the number of changes of sign in the sequence $A_0, A_1, \cdots, A_{n-1}, H(u, s)$, evaluated at K_i, will be exactly k_i. It follows that k_i is the index of K_i.

We wish finally to show that for further suitable restrictions on the constant s_1 the function Ψ has no critical points on the subdomain of Σ_1 for which $0 < s < s_1$ other than those lying on the domains N_i^*. Let B' denote the closure of the set of points of B exclusive of the points on the neighborhoods N_i. With each point Q on B' let there be associated an arbitrarily small open neighborhood U of Q on B admissibly represented by parameters (u). Let U^* denote the domain of points (u, s) for which (u) is on U and $0 \leqq s < s_1$.

If Q is not a critical point of the boundary function φ at least one of the partial derivatives ψ_{uh} does not vanish at Q. The derivatives ψ_{uh} are independent of s_0 and s_1. Accordingly the neighborhood U of Q and the constant s_1 can be chosen so small that at least one of the derivatives ψ_{uh} does not vanish on the corresponding domain U^*.

If Q is a critical point of φ on B' it follows from Lemma 8 that $\psi_s = f_N > 0$ at Q. Now ψ_s depends upon s_0 only through the term Ms^2/s_0^2. Hence a diminution of s_0 merely increases the value of ψ_s. Accordingly the neighborhood U of Q and the constant s_1 can be chosen so small that ψ_s will not vanish on the corresponding domain U^*.

It now follows from the Heine-Borel theorem that B' can be covered with a finite number of open neighborhoods.

$$U_j \qquad (j = 1, \cdots, \nu)$$

for each of which a choice of s_1 can be made such that on the corresponding domains U_j^* at least one of the partial derivatives of ψ does not vanish. The function Ψ has no critical points on these domains U_j^*.

We make a final choice of s_1 as the least of the values s_1 used in defining the respective domains

$$N_1^*, \cdots, N_\mu^*; \qquad U_1^*, \cdots, U_\nu^*.$$

Observe that these domains cover the subdomain of Σ_1 for which $0 < s < s_1$.

The proof of Lemma 9 is complete.

23. **Relations between the critical points of f and the connectivities of $\bar{\Sigma}$.** The preceding function $\Psi(x)$ is of class C'' on Σ_1. The region Σ_0 is a subregion of Σ_1 bounded by B_0. The manifold B_0 is of class C'' and in general not of class C'''. Were B_0 of class C''' the function $\Psi(x)$ would satisfy the boundary conditions β relative to Σ_0 and we could apply Theorem 8 to $\Psi(x)$. To meet this difficulty we introduce the following lemma.

LEMMA 10. *There exists a one-to-one transformation T^* of class C'' which carries Σ_1 into a sub-domain Σ^* and carries Σ_0 into Σ.*

Let the positive constant s_1 have the value chosen at the end of the preceding section. Let $\zeta(s)$ be any real, single-valued function of s which is of class C'' for $-s_1 < s \leqq s_1$ and which satisfies the following conditions:

$$\begin{aligned} &\zeta(s) \equiv s, && -s_1 < s \leqq -s_0, \\ &\zeta(s_0) = 0, \\ &\zeta(s_1) \leqq s_1, \\ &\zeta'(s) > 0, && -s_1 < s \leqq s_1. \end{aligned} \tag{23.1}$$

One can readily establish the existence of such a function. Set $\zeta(s_1) = s^*$. Let Σ^* denote the points of Σ together with the points neighboring B for which $0 \leqq s < s^*$.

On each normal to B the equation $s' = \zeta(s)$ defines a transformation of the points (u, s) for which $-s_1 < s \leqq s_1$ into the points (u, s') for which $-s_1 < s' \leqq s^*$. The transformation thus defined neighboring B is one-to-one and of class C''. Let it be denoted by T. Let T^* be the transformation of Σ_1 into Σ^* in which points of Σ_1 for which $-s_1 < s < s_1$ are transformed under T and in which the remaining points of Σ_1 are transformed by the identity. The transformation T^* is one-to-one and of class C''. Under T^* the region Σ_0 is transformed into the region Σ.

The proof of the lemma is complete.

The relations between the critical points of f and the connectivity numbers of $\bar{\Sigma}$ under general boundary conditions are given in the following theorem.

THEOREM 10. *Let M_i $(i = 0, 1, \cdots, n)$ be the total number of critical points of index i of f on Σ and of φ on the negative boundary of Σ, and let R_j $(j = 0, 1, \cdots, n)$ be the j^{th} connectivity number of $\bar{\Sigma}$. Then under the boundary conditions γ the relations between the numbers M_i and R_j are the same as those in Theorem* 6.

Let $\Psi(x)$ be the function defined in §21. Let the point (x) on Σ^* and the point (x') on Σ_1 correspond under the transformation T^*. On Σ^* we define a function $L(x)$ by the identity:

$$L(x) \equiv \Psi(x').$$

It follows from the properties of T^* and Ψ that $L(x)$ is of class C'' and non-degenerate. Recalling that $\Psi_N > 0$ on B_0 and that $\zeta'(s) > 0$ in (23.1) we see that $L_N > 0$ on B. Each critical point of $f(x)$ is a critical point of $L(x)$ and as a critical point of $L(x)$ has the same index. In addition to these critical points $L(x)$ has one and only one critical point corresponding to each critical point of the boundary function φ on the negative boundary of Σ. The latter critical points lie on the neighborhood of B for which $-s_1 < s < 0$. Each of these critical points has an index equal to that of the corresponding critical point of the boundary function φ.

We see then that $L(x)$ satisfies the boundary conditions β relative to Σ. On applying Theorem 8 to $L(x)$ we obtain Theorem 10.

Bibliography

1. ALEXANDER, J. W., Combinatorial analysis situs, Transactions of the American Mathematical Society, 28 (1926), 301–329.
2. BROWN, A. B., Relations between the critical points and curves of a real analytic function of two independent variables, Annals of Mathematics, 31 (1930), 449–456.
3. BROWN, A. B., Critical sets of an arbitrary real analytic function of n variables, Annals of Mathematics, 32 (1931), 512–520.
4. CAIRNS, S. S., On the cellular sub-division of n-dimensional regions, Annals of Mathematics, 33 (1932), 671–680.
5. DICKSON, L. E., Modern algebraic theories, New York, 1926.
6. JOHN, F., Über die Vollständigkeit der Relationen von Morse für die Anzahlen kritischer Punkte, Mathematische Annalen, 107 (1933), 1–14.
7. v. KERÉKJÁRTÓ, B., Vorlesungen über Topologie I, Berlin, 1923.
8. LEFSCHETZ, S., Topology, American Mathematical Society Colloquium Publications, New York, 1930.
9. MORSE, M., Relations between the critical points of a real function of n independent variables, Transactions of the American Mathematical Society, 27 (1925), 345–396.
10. MORSE, M., The analysis and analysis-situs of regular n-spreads in $(n + r)$-space, Proceedings of the National Academy of Sciences, 13 (1927), 813–817.
11. MORSE, M., The critical points of a function of n variables, Transactions of the American Mathematical Society, 33 (1931), 72–91.
12. MORSE, M., The calculus of variations in the large, American Mathematical Society Colloquium Publications, New York, 1934.
13. POINCARÉ, H., Sur les courbes définies par les équations différentielles, Journal de Mathématiques, série 3, 7 (1881), 375–421.
14. TUCKER, A. W., An abstract approach to manifolds, Annals of Mathematics, 34 (1933), 191–243.
15. VEBLEN, O., Analysis situs, 2nd edition, American Mathematical Society Colloquium Publications, New York, 1931.
16. WHYBURN, W. M., Non-isolated critical points of functions, Bulletin of the American Mathematical Society, 35 (1929), 701–708.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

A THEOREM ON FIXED POINTS FOR PERIODIC TRANSFORMATIONS

BY P. A. SMITH

(Received April 1, 1934)

THEOREM. *Let K be a point set in a cartesian S_m and Λ a topological transformation of K into itself of finite period p. If every continuous single-valued image in K of every sphere of dimension $\leqq pm - m - 1$ is deformable in K to a point, then Λ leaves fixed at least one point.*

The proof will be based on a use of Brouwer's looping coefficients and an exploitation of the symmetry[1] arising from the periodicity of Λ.

1. Let Π be an orthogonal linear transformation of period p operating in a cartesian S_n and leaving fixed only the origin O. We shall denote the transforms of a set A in S_n under successive powers of Π by $A^{(1)}, A^{(2)}, \cdots, A^{(p-1)}$.

Let U_{n-1} be a definitely chosen $(n-1)$-sphere in S_n with center at O. Since Π is linear and transforms U_{n-1} into itself, there is no difficulty in defining a simplicial subdivision U of U_{n-1} which will be preserved under Π; and since Π leaves fixed no point of U_{n-1}, we may assume further that no cell of U is transformed into itself by Π. It is clear that the k-cells of such a subdivision can be so named that their closures will be represented without repetition by

$$e_k^\lambda;\ (e_k^\lambda)^{(1)};\ \cdots;\ (e_k^\lambda)^{(p-1)} \qquad (\lambda = 1; 2, \cdots \lambda_k).$$

Let b_{k-1}^λ be the boundary[2] of e_k^λ and let

$$U_k = \Sigma e_k^\lambda + \Sigma(e_k^\lambda)^{(1)} + \cdots + \Sigma(e_k^\lambda)^{(p-1)} \qquad (k = 0, \cdots, n-2).$$

U_k is thus the sum of the closures of all the k-cells of U; hence $U_0 \subset U_1 \subset \cdots \subset U_{n-1}$. Since e_{k+1}^λ and $(e_{k+1}^\mu)^{(i)}$ have at most boundary points in common, the sets

$$\Sigma e_{k+1} - \Sigma b_k, \qquad \Sigma(e_{k+1})^{(1)} - \Sigma(b_k)^{(1)}, \cdots, \Sigma(e_{k+1})^{(p-1)} - \Sigma(b_k)^{(p-1)}$$

are non-overlapping; their sum is $U_{k+1} - U_k$. Hence if w is an arbitrary point of $U_{k+1} - U_k$, there is a unique point u in $\Sigma e_{k+1} - \Sigma b_k$ and a unique integer i, $(0 \leqq i \leqq p-1)$ such that $w = u^{(i)}$.

2. Let TU_{n-1} consistently denote a continuous single-valued image (or merely, an image) of U_{n-1} in $S_n - O$. If $A \subseteq U_{n-1}$, TA is the image *induced by T, acting only on A*. If SA is a second image of A, we shall write $TA \equiv SA$ if $Ta = Sa$ for every $a \in A$; also, $TA \equiv A$ if $Ta = a$, $a \in A$.

[1] In this respect our methods are somewhat similar to those used by Newman in a paper entitled, A theorem on periodic transformations of spaces, Quart. Journ., vol. 2 (1931), 1–8.

[2] That is, the closure of the boundary of the cell of which e_k^λ is the closure.

We shall denote by $T_tA (0 \leqq t \leqq 1)$ a deformation of TA depending continuously on a parameter t in the interval [0, 1] and such that $T_0A \equiv TA$. It will be understood always that the successive images of a deformation are to lie in $S_n - O$. The relation $T_tA \equiv S_tA$ will mean: $T_ta = S_ta$ for every $t \in [0, 1]$ and $a \in A$.

If $Tu^{(i)} = (Tu)^{(i)}$ for $i = 1, \cdots, p-1$ and every $u \in U_{n-1}$, we shall say that TU_{n-1} is *invariant*; to say that T_tU_{n-1} is invariant will mean that all its successive images are invariant.

In the following lemmas it will often be convenient to denote TA by $T^{[k]}A$, say; we then agree that the deformations represented by the symbols $T_t^{[k+1]}A$ and $T_t^{[k]}A$ are not necessarily related, except of course that $T_0^{[k+1]}A \equiv T_0^{[k]}A \equiv TA$.

3. Let e_{k+1}, $(0 \leqq k \leqq n-2)$ be an arbitrary closed $(k+1)$-cell of the set e_{k+1}^{λ}, and b_k its boundary; then for every deformation $T_t^{[k]}b_k$ of $T^{[k]}b_k$ there exists a $T_t^{[k+1]}e_{k+1}$ such that $T_t^{[k+1]}b_k \equiv T_t^{[k]}b_k$.

PROOF. Think of e_{k+1} as composed of a family b^{δ} $(0 \leqq \delta \leqq 2)$ of "concentric" k-manifolds homeomorphic to k-spheres and depending continuously on δ in such a way that b^0 is a point and $b^2 = b_k$. Let $S\mu\ b^{\delta}$ $(0 < \mu, \delta \leqq 2)$ be the transformation which carries an arbitrary point of b^{δ} into the similarly located point of b^{μ}; in addition, let $S_0\, b^0 = b^0$. The required $T_t^{[k+1]}e_{k+1}$ can be defined by

$$T_t^{[k+1]}b^{\delta} \equiv \begin{cases} S_{f(\delta)}\, b^{\delta} & (0 \leqq \delta \leqq 2-t; \quad f(0) = 0, \quad f(2-t) = 2) \\ T_{g(\lambda)}^{[k]}\,(S_2\, b^{\delta}) & (2-t \leqq \delta \leqq 2; \quad g(2-t) = 0, \quad g(2) = t) \end{cases}$$

where f and g are linear. It will be seen that since the successive images of $T_t^{[k]}b_k$ lie in $S_n - O$, so do those of $T_t^{[k+1]}e_{k+1}$.

4. Let $S\,\Sigma\,e_{k+1}$ $(0 \leqq k \leqq n-3)$ be an image of $\Sigma\,e_{k+1}$ such that $S\,\Sigma\,b_k \equiv \Sigma\,b_k$; there is an $S_t\,\Sigma\,e_{k+1}$ such that $S_t\,\Sigma\,b_k \equiv \Sigma\,b_k$ and $S_1\,\Sigma\,e_{k+1} \equiv \Sigma\,e_{k+1}$.

PROOF. Denoting the i^{th} coordinate of a point by a subscript, define $S_t\,\Sigma\,e_{k+1}$ by

$$(S_t\,u)_i = t(u_i - (Su)_i) + (Su)_i \qquad (u \in \Sigma\,e_{k+1},\ i = 1, \cdots, n).$$

The points of $\Sigma\,b_k$ remain fixed as desired, since $Su = u$ for $u \in \Sigma\,b_k$. Moreover, $S_1\,\Sigma\,e_{k+1} \equiv \Sigma\,e_{k+1}$. There is of course no guarantee that the successive images of $S_t\,\Sigma\,e_{k+1}$ all lie in $S_n - O$. It may in fact happen that the manifold σ of linear segments along which the deformation takes place, contains O. However, by assumption neither of the end images $S\,\Sigma\,e_{k+1}$, $\Sigma\,e_{k+1}$ contains O; moreover, σ is an image of a complex of $k+2$ dimensions. Since $k+2 \leqq n-1$, σ can be modified near O without disturbing $S\,\Sigma\,e_{k+1}$ or $\Sigma\,e_{k+1}$, and in such a way that it will no longer contain O. The deformation across the modified manifold is of the required type.

5. Suppose TU_{n-1} is invariant. Then if there is an invariant $T_t^{[k]}\, U_k$ $(0 \leqq k \leqq n-3)$ such that $T_1^{[k]}\, U_k \equiv U_k$, the same is true for the dimension $k+1$.

PROOF. By §3, we have for $\lambda = 1, \cdots, \lambda_k$ a $T_t^{[k+1]}\, e_{k+1}^{\lambda}$ such that $T_t^{[k+1]}\, b_k^{\lambda} \equiv T_t^{[k]}\, b_k^{\lambda}$. These deformations taken together define a $T_t^{[k+1]}\, \Sigma\, e_{k+1}$ such that $T_t^{[k+1]}\, \Sigma\, b_k \equiv T^{[k]}\, \Sigma\, b_k$. By a change of parameter, let the range of t for $T_t^{[k+1]}$ be reduced to $[0, \frac{1}{2}]$ and let $S\, \Sigma\, e_{k+1} \equiv T_{\frac{1}{2}}^{[k+1]}\, \Sigma\, e_{k+1}$ so that

$$S\, \Sigma\, b_k \equiv T_{\frac{1}{2}}^{[k+1]}\, \Sigma\, b_k \equiv T_1^{[k]}\, \Sigma\, b_k \equiv \Sigma\, b_k.$$

By §4, there is an $S_t\, \Sigma\, e_{k+1}$ such that $S_t\, \Sigma\, b_k \equiv S\, \Sigma\, b_k$ and $S_1\, \Sigma\, e_{k+1} \equiv \Sigma\, e_{k+1}$. Now define $T_t^{[k+1]}\, \Sigma\, e_{k+1}$ for the interval $[\frac{1}{2}, 1]$ by

$$T_t^{[k+1]}\, \Sigma\, e_{k+1} \equiv S_{2t-1}\, \Sigma\, e_{k+1} \qquad (\tfrac{1}{2} \leqq t \leqq 1).$$

We thus obtain a $T_t^{[k+1]}\, \Sigma\, e_{k+1}$ $(0 \leqq t \leqq 1)$ with the properties:

$$T_1^{[k+1]}\, \Sigma\, e_{k+1} \equiv \Sigma\, e_{k+1}$$

$$T_t^{[k+1]}\, \Sigma\, b_k \equiv T_{2t}^{[k]}\, \Sigma\, b_k \ (0 \leqq t \leqq \tfrac{1}{2}); \qquad \equiv \Sigma\, b_k \ (\tfrac{1}{2} \leqq t \leqq 1).$$

We now extend the definition of $T_t^{[k+1]}$ to the whole of U_k simply by defining

$$T_t^{[k+1]}\, U_k \equiv T_{2t}^{[k]}\, U_k \ (0 \leqq t \leqq \tfrac{1}{2}); \equiv U_k \ (\tfrac{1}{2} \leqq t \leqq 1).$$

Since $T_t^{[k]}\, U_k$ is invariant, so is $T_t^{[k+1]}\, U_k$. Next we extend the definition of $T_t^{[k+1]}$ to the whole of U_{k+1} by defining

$$T_t^{[k+1]}\, u^{(i)} = (T_t^{[k+1]} u)^{(i)} \qquad (i = 1, \cdots, p-1) \tag{1}$$

for $u \in \Sigma\, e_{k+1} - \Sigma\, b_k$. Since every w in $U_{k+1} - U_k$ is uniquely expressible in the form $u^{(i)}$, $(0 \leqq i \leqq p-1,\ u \in \Sigma\, e_{k+1} - \Sigma\, b_k)$ (see §1) (1) defines a unique image for every w in $U_{k+1} - U_k$. Hence $T_t^{[k+1]}\, w$ is uniquely defined for every w in U_{k+1}. Obviously $T_t^{[k+1]}$ is invariant and satisfies the condition $T_t^{[k+1]}\, U_k \equiv T^{[k]}\, U_k$. We still have to verify the continuity of $T_t^{[k+1]}$ for a fixed t.[3] Suppose w is an arbitrary point of $(e_k^{\lambda})^{(i)} = e^{(i)}$ so that $w = u^{(i)}$, $u \in e$. Then

$$T_t^{[k+1]}\, w = T_t^{[k+1]}\, u^{(i)} = (T_t^{[k+1]}\, u)^{(i)}.$$

The second equality holds for $u \in e - b$ on account of the invariance of $T_t^{[k+1]}\, U_k$; and for $u \in b$ by virtue of (1). Now since $T_t^{[k+1]}\, u$ is a continuous function of u for $u \in e$, so is $(T_t^{[k+1]}\, u)^{(i)}$; hence $T^{[k+1]}\, w$ is continuous for $w \in e^{(i)}$. Since $e^{(i)}$ is an arbitrary closed cell of U_{k+1} the continuity of $T_t^{[k+1]}\, U_{k+1}$ is established.

6. If TU_{n-1} is invariant, there is an invariant $T_t\, U_{n-1}$ such that $T_1\, U_{n-2} \equiv U_{n-2}$.

PROOF. Choose for $\lambda = 1, \cdots, \lambda_0$ a $T_t^{[0]}\, e_0^{\lambda}$ such that $T_1^{[0]}\, e_0^{\lambda} \equiv e_0^{\lambda}$ and let $T_t^{[0]}\, (e_0^{\lambda})^{(i)} = (T_t^{[0]}\, e_0^{\lambda})^{(i)}$ $(i = 1, \cdots, p-1)$. We obtain thus an invariant

[3] Continuity with respect to t is readily verified.

$T_t^{[0]} U_0$ such that $T_1^0 U_0 \equiv U_0$. Now apply §5 $n-2$ successive times and obtain an invariant $T_t^{[n-2]} U_{n-2}$ such that $T_1^{[n-2]} U_{n-2} \equiv U_{n-2}$. By §3, there is a $T_t^{[n-1]} \Sigma e_{n-1}$ such that $T_t^{[n-1]} \Sigma b_{n-2} \equiv T_t^{[n-2]} \Sigma b_{n-2} \equiv \Sigma b_{n-2}$. Extend the definition of $T_t^{[n-1]}$ to the whole of U_{n-2} by defining $T_t^{[n-1]} U_{n-2} \equiv U_{n-2}$ then to the whole of U_{n-1} by defining $T_t^{[n-1]} u^{(i)} = (T_t^{[n-1]} u)^{(i)}$ $(i = 1, \cdots, p-1)$ for $u \in \Sigma e_{n-1} - \Sigma b_n$. There is thus defined, by the argument of §5 a $T_t^{[n-1]} U_{n-1}$ with the desired properties.

7. To an image TU_{n-1} there can be attached an integer which denotes the number of times that TU_{n-1} "loops" O. We shall review briefly the definition and properties of this looping coefficient.[4]

Let definite orientations be chosen for U_{n-1} and S_n and let E^λ $(\lambda = 1, \cdots \alpha)$ be the $(n-1)$-cells, positively oriented, of a simplicial subdivision V of U_{n-1}. If the vertices of E^λ are $a_\lambda, b_\lambda, \cdots, c_\lambda$, named in an order corresponding to the orientation of E^λ, we write $E^\lambda = a_\lambda b_\lambda \cdots c_\lambda$. Let the oriented *flat* simplex F^λ (possibly degenerate) determined by the same vertices named in the same order, be denoted by $\{a_\lambda b_\lambda \cdots c_\lambda\}$. If the subdivision V is sufficiently fine, O will not be contained in any F^λ; hence F^λ subtends at O a solid angle of measure $[F^\lambda]$; we agree to include a $+$ or $-$ in the symbol $[F^\lambda]$ according as the orientation of $O \cdot F^\lambda$ is positive or negative.[5] The cycle ΣF^λ is a simplicial approximation to U_{n-1} and if V is fine enough, $\Sigma[F^\lambda] = \pm\mu$ where μ is the solid angle subtended by U_{n-1},—that is, the $(n-1)$-volume of an $(n-1)$-sphere of unit radius. In the same way the simplex $F_T^\lambda = \{Ta_\lambda\, Tb_\lambda \cdots Tc_\lambda\}$ subtends a solid angle of measure $[F_T^\lambda]$ and the quantity $\Sigma[F_T^\lambda]$ is independent of the subdivision V (provided only that its mesh is sufficiently small) and hence can be denoted by $[TU_{n-1}]$. Its value is $N\mu$ where $|N|$,—the looping coefficient mentioned above,—is zero or an integer. If $T_t U_{n-1}$ fails to meet O for any t,—and this is the only case considered, $[T_t U_{n-1}]$ is a continuous function of t, and since it only assumes discreet values, it is constant.

8. We shall now prove that if TU_{n-1} is invariant, then $[TU_{n-1}] \neq 0$. First, we may assume that $TU_{n-2} \equiv U_{n-2}$; for by §6, this situation can be obtained by a deformation, and as we have just remarked, the value of $[TU_{n-1}]$ is not thereby changed. We shall now take for V a subdivision of U so chosen, that V, like U, is preserved under II. Let the $(n-1)$-cells of V be so named that those in Σe_{n-1}, positively oriented, are $E^1, \cdots, E^h$. Then the totality of positively oriented $(n-1)$-cells are

$$E^\lambda;\ (\eta_1 E^\lambda)^{(1)};\ \cdots;\ (\eta_{p-1} E^\lambda)^{(p-1)} \qquad (\lambda = 1, \cdots, h)$$

where η_i is $+1$ or -1 according as the transformation $u \to u^{(i)}$ preserves or reverses the orientation of U_{n-1}. Since the orientation of S_n is influenced by this

[4] For details, see Brouwer, On looping coefficients, Amsterdam proceedings, 15 (1912) 113–122.

[5] $O \cdot F^\lambda$ denotes the oriented simplex $\{Oa_\lambda\, b_\lambda \cdots c_\lambda\}$.

transformation in the same way as that of U_{n-1}, it follows that for each λ and i, $(O\cdot\eta_i F^\lambda)^{(i)}$ is[6] concordant with $O\cdot F^\lambda$. Now $(O\cdot\eta_i F^\lambda)^{(i)}$ and $(O\cdot\eta_i F_T^\lambda)^{(i)}$ are concordant or not according as $O\cdot\eta_i F^\lambda$ and $O\cdot\eta_i F_T^\lambda$ are or are not; hence, as $O\cdot F^\lambda$ and $O\cdot F_T^\lambda$ are or are not. Since $O\cdot F^\lambda$ and $(O\cdot\eta_i F^\lambda)^{(i)}$ are concordant, it follows that $O\cdot F_T^\lambda$ and $(O\cdot\eta_i F_T^\lambda)^{(i)} = O\cdot(\eta_i F_T^\lambda)^{(i)}$ are concordant; hence $[F_T^\lambda]$ and $[\eta_i F_T^\lambda]^{(i)}$ are of same sign. Since Π is orthogonal, the measures of the angles subtended by F_T^λ and $(\eta_i F_T^\lambda)^{(i)}$ are numerically equal; it follows that

$$[F_T^\lambda] = [(\eta_i F_T^\lambda)^{(i)}] \qquad (\lambda = 1, \cdots, h;\ i = 1, \cdots, p-1). \tag{2}$$

The flat cell determined by $(\eta_i E^\lambda)^{(i)}$ is $\eta_i\{a_\lambda b_\lambda \cdots c_\lambda\}^{(i)}$. Hence that determined by $(\eta_i E^\lambda)^{(i)}$ and T is $\eta_i\{Ta_\lambda \cdots Tc_\lambda\}^{(i)}$, which can be written $\eta_i\{(Ta_\lambda)^{(i)} \cdots (Tc_\lambda)^{(i)}\}$ since T is invariant; this cell is precisely $\eta_i(F_T^\lambda)^{(i)} = (\eta_i F_T^\lambda)^{(i)}$. Thus the flat cells used in evaluating TU_{n-1} are

$$F_T^\lambda;\ (\eta_1 F_T^\lambda)^{(1)};\ \cdots\ (\eta_{p-1} F_T^\lambda)^{(p-1)} \qquad (\lambda = 1, \cdots, h),$$

and accordingly,

$$[TU_{n-1}] = \Sigma[F_T^\lambda] + \Sigma[(\eta_1 F_T^\lambda)^{(1)}] + \cdots + \Sigma[(\eta_{p-1} F_T^\lambda)^{(p-1)}].$$

Hence by (2)

$$[TU_{n-1}] = p\Sigma[F_T^\lambda] = p\Sigma[(\eta_1 F_T^\lambda)^1] = \cdots = p\Sigma[(\eta_{p-1} F_T^\lambda)^{(p-1)}]. \tag{3}$$

If for the moment we suppose that $TU_{n-1} \equiv U_{n-1}$, we obtain

$$\pm\mu = p\Sigma[F^\lambda] = p\Sigma[\eta_1 F^\lambda)^{(1)}] = \cdots = p\Sigma[(\eta_1 F^\lambda)^{(p-1)}]. \tag{4}$$

Consider now the image SU_{n-1} defined by

$$Su = Tu \quad \text{for} \quad u \in \Sigma e_{n-1}^\lambda; \qquad = u \text{ for } u \in \Sigma(e_{n-1}^\lambda)^{(1)} + \cdots + \Sigma(e_{n-1}^\lambda)^{(p-1)}.$$

The single-valuedness and continuity of S follow from the relation

$$TU_{n-2} \equiv U_{n-2}.$$

Clearly

$$[SU_{n-1}] = \Sigma[F_T^\lambda] + \Sigma[(\eta_1 F^\lambda)^{(1)}] + \cdots + \Sigma[(\eta_{p-1} F^\lambda)^{p-1}]$$

$$= \Sigma[F_T^\lambda] \pm \frac{p-1}{p}\,\mu \qquad \text{by} \qquad (4).$$

But since $[SU_{n-1}] = \mathrm{M}$ where $|\,\mathrm{M}\,|$ is zero or an integer, it follows that $\Sigma[F_T^\lambda] \neq 0$. Hence by (3), $[TU_{n-1}] \neq 0$.

9. Let

$$x \times y \times \cdots \times z = x_1, \cdots x_m, y_1, \cdots, y_m, \cdots, z_1, \cdots z_m (p \text{ letters})$$

[6] $\{Oab \cdots c\}^{(i)}$ is the simplex $\{O^{(i)} a^{(i)} \cdots c^{(i)}\} = O\cdot\{a^{(i)} \cdots c^{(i)}\}$.

be an arbitrary point of a cartesian S_{pm} and let $I_m = \{x \times x \times \cdots \times x\}$.[7] Let $n = pm - m$ and take S_n to be a flat subspace of S_{pm}, orthogonal to I_m and intersecting I_m at the single point O. The transformation

$$(5) \qquad (x \times y \times \cdots \times z)' = (y \times \cdots \times z \times x)$$

of S_{pm} is of period p; it is also linear and orthogonal and leaves fixed only the points of I_m; consequently it transforms S_n into itself linearly and orthogonally, leaving fixed only the point O. If $P(x \times y \times \cdots \times z)$ denotes the orthogonal projection of $x \times y \times \cdots \times z$ on S_n, then

$$P(x \times y \times \cdots \times z)' = (P(x \times y \times \cdots \times z))'$$

on account of the orthogonality of (5).

Consider now a topological transformation of period p:

$$(\Lambda) \qquad x \to x^{(1)} \to x^{(2)} \to \cdots \to x^{(p-1)} \to x; \ (x, x^{(1)}, \cdots, x^{(p-1)} \in K \subseteq S_m)$$

where K satisfies the conditions of the theorem. The space

$$\Gamma = \{x \times x^{(1)} \times \cdots \times x^{(p-1)}\} \qquad (x \in K)$$

is homeomorphic to K, and is transformed into itself by (5) since

$$(x \times x^{(1)} \times \cdots \times x^{(p-1)})' = x^{(1)} \times \cdots \times x^{(p-1)} \times x$$

$$= x^{(1)} \times (x^{(1)})^{(1)} \times \cdots \times (x^{(1)})^{(p-1)} = \text{a point of } \Gamma.$$

The existence of a fixed point for Λ will be established if we show that Γ meets I_m; for, a point $x \times x^{(1)} \times \cdots \times x^{(p-1)}$ of Γ is on I_m only if $x = x^{(1)}$.

We shall construct in Γ an image τU_{n-1} such that

$$(6) \qquad \tau u^{(i)} = (\tau u)^{(i)} \qquad (i = 1, \cdots, p-1).$$

for every u in U_{n-1}. Let $\tau^{[0]} e_0^\lambda$ $(\lambda = 1, \cdots \lambda_0)$ be arbitrary distinct points in Γ; then let

$$\tau^{[0]} (e_0^\lambda)^{(i)} = (\tau^{[0]} e_0^\lambda)^{(i)} \qquad (\lambda = 1, \cdots, \lambda_0; i = 1, \cdots, p-1).$$

We thus obtain a $\tau^0 U_0$ invariant in the sense (6). Since Γ is homeomorphic to K, every image in Γ of a k-sphere, $k \leqq pm - m - 1 = n - 1$, is deformable in Γ to a point, hence is the boundary of a singular $(k+1)$-cell in Γ. Hence in particular there is for $\lambda = 1, \cdots \lambda_1$ a $\tau^{[1]} e_1^\lambda$ such that $\tau^{[1]} b_0^\lambda \equiv \tau^{[0]} b_0^\lambda$. Now define

$$\tau^{[1]} (e_1^\lambda)^{(i)} = (\tau^{[1]} e_1^\lambda)^{(i)} \qquad \text{for} \qquad \lambda = 1, \cdots \lambda_1; i = 1, \cdots, p-1$$

obtaining thus a $\tau^{[1]} U_1$ invariant in the sense (6). It is clear that this is the first step of a process which can be continued until we obtain the desired τU_{n-1} $(\equiv \tau^{[n-1]} U_{n-1})$.

---

[7] In this § we use $\{\}$ to denote the totality of points of the enclosed type.

Now let $TU_{n-1} \equiv P(\tau U_{n-1})$; TU_{n-1} is invariant since

$$Tu^{(i)} = P(\tau u^{(i)}) = P(\tau u)^{(i)} = (P(\tau u))^{(i)} = (Tu)^{(i)} \qquad (u \in U_{n-1}).$$

Let $\tau_t U_{n-1}$ be a deformation in Γ such that $\tau_1 U_{n-1}$ is a point. Let $T_t U_{n-1} \equiv P(\tau_t U_{n-1})$. Then for some $t_0 (0 \leqq t_0 \leqq 1)$ $T_{t_0} U_{n-1}$ contains O. For if not, then $[T_t U_{n-1}] = [TU_{n-1}]$ $(0 \leqq t \leqq 1)$ by §7, and $[TU_{n-1}] \neq 0$ by §8; but $[T_1 U_{n-1}] = 0$ since $T_1 U_{n-1}$ is a point. Since the only points of S_{pm} which project onto O are those of I_m, it follows that I_m is intersected by $\tau_{t_0} U_{n-1}$ and hence by Γ which finally proves our theorem.

REMARK. An example shows that the integer $pm - m - 1$ can not be replaced by a smaller one. Let K be an $(m-1)$-sphere in S_m, and let Λ be a reflection of K across its center, so that $p = 2$. Every image in K of every sphere of dimension $\leqq m - 2 = pm - m - 2$ is deformable in K to a point, but Λ admits no fixed points.

APPLICATION. Suppose the points of K are the elements of a group with continuous multiplication. If the group K admits an element of finite period p, then for some integer $k \leqq pm - m - 1$ there is in K an image of a k-sphere which is not deformable to a point. For suppose that $a^p = e =$ the identity. Then the transformation

$$x \to ax \qquad (x \in S)$$

is of period p and if our assertion is false, it admits an invariant point say x_0, so that $x_0 = ax_0$; this, however, would imply $a = e$.

COLUMBIA UNIVERSITY.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

ON THE TRIANGULATION OF REGULAR LOCI[1]

BY STEWART S. CAIRNS

(Received June 3, 1932)

1. **Introduction.** In an earlier paper,[2] the writer presented a method for dividing a closed n-dimensional region with certain regularity restrictions on its boundary into the cells of a topological complex. Triangulation theorems for analytic varieties have been discussed by van der Waerden,[3] Lefschetz,[4] Koopman and Brown.[5] In the present paper, new methods are employed to extend the known results to a large class of loci, subject to light regularity conditions.

2. **Regular manifolds and loci.** Let $(y) \equiv (y_1, \cdots, y_n)$ be coordinates in an n-space, and let $(u) \equiv (u_1, \cdots, u_r)$ be rectangular cartesian coordinates in an auxiliary euclidean r-space ($r \leqq n$). An r-cell, open or closed,[6] in the n-space will be called *regular* if it can be defined analytically by a homeomorphism

$$y_i = f_i(u) \qquad (i = 1, \cdots, n) \tag{2.1}$$

with some open or closed convex polyhedral r-cell, s_r, where (1) the f's are of class[7] C' on s_r and (2) the matrix $\left(\frac{\partial f_i}{\partial u_j}\right)$ is of rank r on s_r. Such a homeomorphism will be called a *defining homeomorphism* of the cell.

An *interior point* of an r-dimensional set, S_r, will mean a point one of whose neighborhoods on S_r is an r-cell. A *boundary point* will mean a point, P, one of whose neighborhoods on S_r is a closed r-cell with P on its boundary. An interior or boundary point will be called *regular* if one of its neighborhoods on S_r is an open or closed regular r-cell.

By an *r-manifold* we mean a compact connected r-dimensional point set made up of interior and boundary points.

---

[1] This paper is a generalization of part of the writer's thesis, *The cellular subdivision and approximation of regular spreads* (Harvard, 1931), written under Professor Marston Morse. The results appear in three abstracts: Proc. Nat. Acad. Sci., 16 (1930) pp. 488–491; Bull. Amer. Math. Soc., 38 (1932) p. 807, and 39(1933) p. 349.

[2] *On the cellular subdivision of n-dimensional regions*, these Annals, 33 (1932) pp. 671–680.

[3] *Topologische Begründung des Kalküls der abzählenden Geometrie, Anhang I*, Math. Ann. 102 (1929) pp. 360–361.

[4] *Topology* (1930), Chapter VIII. For further details, see S. Lefschetz and J. H C. Whitehead, *On analytic complexes*, Trans. Amer. Math. Soc., 35 (1933) pp. 510–517.

[5] *On the covering of analytic loci by complexes*, Trans. Amer. Math. Soc., 34 (1932) pp. 231–251.

[6] By a *closed r-cell* we mean the closure of an r-cell.

[7] A function is said to be of class C′ on a point set if it has continuous first partial derivatives at interior points and if these derivatives have continuous limiting values at boundary points.

A *regular* or *completely regular* 0-*manifold* means a point. A *regular* 0-*locus* will mean a finite set of points. To obtain a recurrent definition, assume the terms *completely regular i-manifold* and *regular i-locus*[8] defined for $i = 0, \cdots, r-1$. An r-manifold will be called *completely regular* if all its points are regular and either (1) it has no boundary points or (2) its boundary consists of a finite number of distinct regular unbounded $(r-1)$-manifolds. A *regular r-locus*, L_r, will mean a finite set of completely regular i-manifolds $(i = 0, \cdots, r)$, called *elements of* L_r, any two of which (M_i, M_j) are so related that if an interior point of M_i is on M_j, then M_i is an element of the boundary of M_j. It should be noted that M_j contains its boundary and that every element of its boundary is accordingly an element of L_r.

3. **The three sets of covering cells.** Let L_r be any regular r-locus in the n-space of the coordinates (y), and let M_r denote any r-dimensional element of L_r. Our first object will be to triangulate M_r.

(A) It is possible to select on M_r three sets of closed regular r-cells $(\alpha) \equiv (\alpha^1, \cdots, \alpha^\nu)$, $(\beta) \equiv (\beta^1, \cdots, \beta^\nu)$, and $(\gamma) \equiv (\gamma^1, \cdots, \gamma^\nu)$ with the following properties:

(1) $\alpha^i \supset \beta^i \supset \gamma^i$ $(i = 1, \cdots, \nu)$.

(2) Every point on M_r has one of the cells (γ) for a neighborhood on M_r.

(3) γ^i contains no limit point of $(M_r - \beta^i)$ and β^i contains no limit point of $(M_r - \alpha^i)$.

Any point P on M_r has some closed regular r-cell, α, for a neighborhood on M_r. Under a defining homeomorphism H (§2) α corresponds to a closed convex polyhedral r-cell, a. There is no difficulty in selecting on a two closed r-cells, b and c, whose images, β and γ, under H are such that parts (1) and (3) of (A), read without the superscripts, are satisfied and that γ is a neighborhood of P on M_r. It remains only to apply the Heine-Borel theorem.

(B) Let H^i denote a defining homeomorphism (§2) of α^i, and let (a^i, b^i, c^i) denote the r-cells to which $(\alpha^i, \beta^i, \gamma^i)$ correspond under H^i $(i = 1, \cdots, \nu)$. Then the distance[9] from γ^i to the closure of $(M_r - \beta^i)$ is positive, and so is the distance from β^i to the closure of $(M_r - \alpha^i)$. We will denote by d half the shortest of these 2ν distances.

4. **The homeomorphisms** H^{ik}. In a euclidean r-space with coordinates $(u) \equiv (u_1, \cdots, u_r)$ an arbitrary m-plane, L_m, can be parametrically defined as follows:

$$u_i = \sum_{j=1}^{m} a_{ij} t_j + u_i^0 \qquad (i = 1, \cdots, r). \tag{4.1}$$

[8] A *regular i-manifold*, for any value of i, will mean a regular i-locus which is a manifold.

[9] We occasionally write as if α^i and a^i were identical, each point coinciding with its image under H^i. This permits us to use metric terms on α^i and, in §6, to speak of $H_m^{i_1 \cdots i_p}$-cells on a^i.

Let $\delta_{i_1 \cdots i_m}$ denote the determinant $|a_{ij}|$ $(i = i_1, \cdots, i_m;\ j = 1, \cdots, m)$, where $(i_1, \cdots, i_m)$ are m numbers in order of increasing magnitude from the set $(1, \cdots, r)$. The determinants $\delta_{i_1 \cdots i_m}$ will be called *direction components* of L_m and the ratios

$$\gamma_{i_1 \cdots i_m} = \frac{\delta_{i_1 \cdots i_m}}{[\Sigma\, \delta^2_{i_1 \cdots i_m}]^{\frac{1}{2}}} \tag{4.2}$$

will be called *direction cosines*, where Σ is a summation over all the δ's. The writer has shown[10] that $\gamma_{i_1 \cdots i_m}$ is numerically the inverse ratio of an m-dimensional volume on L_m to its projection on the coordinate plane of $(y_{i_1}, \cdots, y_{i_m})$.

LEMMA 1. *Let (α^i, α^k) be two overlapping cells of the set (α) [§3(A)]. Consider their defining homeomorphisms* [§3(B)]

$$\begin{aligned} H^i\colon &\quad y_i = f_i(u) &&(i = 1, \cdots, n),\\ H^k\colon &\quad y_i = g_i(v) &&(i = 1, \cdots, n). \end{aligned} \tag{4.3}$$

Given any point P of the product $\alpha^i \cdot \alpha^k$, let $J(f_{i_1} \cdots f_{i_r})$ be any jacobian of r of the f's which is not zero at P. Then $J(g_{i_1} \cdots g_{i_r})$ is not zero at P.

For, let (y) be interpreted as coordinates in a euclidean n-space. Then the jacobians of the f's, figured at P, are a set of direction components of the tangent r-plane to M_r at P, and so are the jacobians of the g's. Our result follows at once.

(A) *Under the conditions of Lemma* 1, *Let A^i and A^k be the images of the product $\alpha^i \cdot \alpha^k$ under H^i and H^k respectively. Consider the homeomorphism H^{ik}*

$$v_i = F_i(u) \qquad (i = 1, \cdots, r) \tag{4.4}$$

between A^i and A^k in which two points correspond if their images on $\alpha^i \cdot \alpha^k$ coincide. The functions F_i are then of class[7] C' on A^i and $J(F_1 \cdots F_r) \neq 0$ on A^i.

This is easily shown[11], making use of the fact that, under the hypotheses of Lemma 1,

$$J(F_1 \cdots F_r) \equiv \frac{J(g_{i_1} \cdots g_{i_r})}{J(f_{i_1} \cdots f_{i_r})}.$$

(B)[11] A regular m-cell (§2), open or closed, in a euclidean r-space has a tangent m-plane at any point, P, on the cell. The direction cosines of the tangent m-plane vary continuously with P.

(C)[11] The image under H^{ik} of a regular m-cell, open or closed, on A^i is a regular m-cell on A^k.

LEMMA 2. *For any value $\epsilon > 0$, there exists a value $\delta > 0$ so small that if σ_m is a regular (open or closed) m-cell on A^i of diameter less than δ on which no direction*

[10] *The direction cosines of a p-space in euclidean n-space*, Amer. Math. Monthly, 39 (1932) pp. 518–523.

[11] Details of the proof are omitted, for a direct analytic verification offers no difficulties.

cosine of the tangent m-plane varies by more than δ, then the image of σ_m under H^{ik} is a regular m-cell on A^k of diameter less than ϵ, on which no direction cosine varies by more than ϵ.

For, let L_m be any m-plane through an interior point P of A^j. Let P' be the image of P under H^{ik} and let L_m' be the m-plane tangent at P' to the image of L_m. Then the coordinates of P' and the direction cosines of L_m' can be expressed as continuous functions of the coordinates of P and the direction cosines of L_m. These functions are[11] defined and continuous even if P is a boundary point of A^j. Hence P and the direction cosines of L_m can vary over a closed set. Our lemma now follows readily.

5. The $H_m^{i_1 \cdots i_p}$-cells. Any point on α^j ($j = 1, \cdots, \nu$) [see §3 (A)] will be called an H_0^j-*cell*. A 1-cell on α^j will be called an H_1^j-*cell* if its image under H^j [§3 (B)] is a line-segment. We will define an $H_m^{i_1 \cdots i_p}$-*cell* by a recurrent process in such a way as to preserve the following properties, which are easily verified in the case of an H_1^j-cell.

(A) *The image under H^k ($k = 1, \cdots, \nu$) [§3 (B)] of an $H_m^{i_1 \cdots i_p}$-cell, σ_m, on α^k [§3 (A)] is a regular simplicial m-cell. The bounding cells of an $H_m^{i_1 \cdots i_p}$-cell are $H_j^{k_1 \cdots k_q}$-cells ($j = 0, \cdots, m-1$) where $(k_1, \cdots, k_q)$ ranges through various subsets of $(i_1, \cdots, i_p)$.*

Our recurrent process introduces the vertices of an $H_m^{i_1 \cdots i_p}$-cell ($p = 1, \cdots, r+1$; $m = p-1, \cdots, r$) one at a time. Let σ_{m-1} denote either an $H_{m-1}^{i_1 \cdots i_{p-1}}$-cell or an $H_{m-1}^{i_1 \cdots i_p}$-cell on α^{i_p}, where $(i_1, \cdots, i_p)$ are distinct, and let P denote any point on[12] $(\alpha^{i_p} - \bar{\sigma}_{m-1})$.

(B) *Suppose the $H_1^{i_p}$-cells joining P to $\bar{\sigma}_{m-1}$ are distinct and none of them is tangent to σ_{m-1} or to one of its bounding cells.*[13] *Then the m-cell σ_m, composed of the $H_1^{i_p}$-cells joining P to σ_{m-1}, will be called an $H_m^{i_1 \cdots i_p}$-cell. By its bounding cells, we will mean the point P, the cell σ_{m-1} and its bounding cells, and every q-cell composed of the $H_1^{i_p}$-cells which join P to a bounding $(q-1)$-cell of σ_{m-1}.*

Before showing that properties (A) are preserved, we draw certain consequences from our recurrent process. The homeomorphism H^i affords a definition of straightness on α^i ($i = 1, \cdots, \nu$). As we build up an $H_m^{i_1 \cdots i_p}$-cell in accordance with (B), we employ, during the first few steps, straightness as defined by H^{i_1}, then straightness as defined by H^{i_2}, and so on.

(C) *Hence an H_m^j-cell is the image under H^j of an m-simplex. In general, an $H_m^{i_1 \cdots i_p}$-cell, σ_m, is uniquely determined by some pair consisting of an $H_j^{i_1 \cdots i_{p-1}}$-cell, σ_j, and an $H_k^{i_p}$-cell, σ_k, where $m = j + k + 1$. The cell σ_m is the set of all points each on one of the $H_{k+1}^{i_p}$-cells determined by σ_k and a point P as P ranges over σ_j.*

---

[12] The symbol for a point set, modified by a bar, denotes the closure of the set.

[13] That is, if (s_{m-1}, Q) correspond under H^{i_p} to (σ_{m-1}, P), no line-segment joining Q to s_{m-1} is in a tangent plane to s_{m-1} or to one of its bounding cells [cf §4(B) and §5(A)].

While we might have used (C) as the basis for a more compact definition of $H_m^{i_1 \cdots i_p}$-cells, the process of introducing one vertex at a time simplifies certain inductive segments in the remainder of the paper.

Our definition obviously preserves the second part of (A). As for the first part, suppose (s_{m-1}, Q) correspond under H^{i_p} to (σ_{m-1}, P). Then s_{m-1} is by hypothesis a regular $(m-1)$-cell. Since the segments joining Q to s_{m-1} are distinct and none of them is tangent to s_{m-1}, these segments form[11] a regular m-cell, s_m, which corresponds under H^{i_p} to σ_m. The image of σ_m under H^k is, by §4 (A), the image of s_m under H^{jk}, where $j = i_p$. Our result now follows from §4 (C).

(D) Let (P) be a set of $(r+1)$ points on M_r, and let $(\beta^{j_1}, \cdots, \beta^{j_q})$ denote all the cells of the set (β) [§3 (A)] each of which contains one or more of the points (P). We will call (P) *admissible* if

(1) (P) is on α^i $(i = j_1, \cdots, j_q)$ [§3(A)] and its image under H^i is of diameter less than d [§3(B)].

(2) The points (P) determine an H_r^i-cell $(i = j_1, \cdots, j_q)$ [cf (C)].

(3) Every pair consisting of an $H_j^{i_1 \cdots i_{p-1}}$-cell and an $H_k^{i_p}$-cell with no common vertices and with all their vertices in the set (P) determines an $H_{j+k+l}^{i_1 \cdots i_p}$-cell [cf(C)], provided $(i_1, \cdots, i_p)$ are p distinct numbers from the set $(j_1, \cdots, j_q)$.

6. **A sufficient condition for admissibility.** Let ϑ be a positive acute angle. A set (P) of $r+1$ points in euclidean r-space will be called a *ϑ-set* if it determines an r-simplex for which every angle between an $(r-1)$-dimensional face and an edge incident with the opposite vertex exceeds ϑ. The object of this section is to establish the following theorem.

THEOREM. *Let (P) be a set of $(r+1)$ points on a cell, α^j, of the set (α) [§3(A)], where at least one of the points (P) is on β^j. For any positive acute angle ϑ a distance $\rho(\vartheta) > 0$ exists so small that if the image (Q) of (P) under H^j [§3(B)] is a ϑ-set of diameter less than $\rho(\vartheta)$, then (P) is admissible* [§5(D)].

Let $\rho_1 > 0$ be so small that, independently of j and k, a point set on A^j [§4(A)] of diameter less than ρ_1 has for its image under H^{jk} a set of diameter less than d [§3(B)]. Then if (Q) is of diameter less than ρ_1, (P) will satisfy part (1) of the definition of admissibility [§5(D)].

LEMMA 1. *Given the angle ϑ, there exists an angle $\vartheta' > 0$ and a distance $\rho_2 > 0$, so small that, independently of j and k, if (Q) is a ϑ-set of diameter less than ρ_2 on A^j [§4(A)], then the image (R) of (Q) under H^{jk} is a ϑ'-set.*

In the neighborhood of any point (u^0) of A^j, the linear correspondence [cf (4.4)]

$$v_i = \sum_{j=1}^{r} \left(\frac{\partial F_i}{\partial u_j}\right)^0 u_j + u_i^0 \qquad (i = 1, \cdots, r) \tag{6.1}$$

is a first degree approximation to H^{jk}. If (R′) denote the image of (Q) under (6.1), then the ratio of the r-simplex determined by (R') to the one determined by (Q) is the value of $J(F_1 \cdots F_r)$ at (u^0). But if (u^0) is taken as one of the points (Q), then (R') can be made an arbitrarily close approximation to (R) by

requiring that (Q) be a sufficiently small ϑ-set. The proof is easy to complete.[11]

If (Q) is a ϑ-set of diameter less than both ρ_1 and ρ_2, then (P) will satisfy parts (1) *and* (2) *of the definition of admissibility* [§5(D)]. *It remains only to deal with part* (3) *of that definition.*

Let σ_m denote the image under H^j of any $H_m^{i_1 \cdots i_p}$-cell whose vertices belong to the set (P) (see §5(D) and the above theorem for notation). Let s_m be the m-simplex determined by the vertices of σ_m, and let s_1 be the length of the shortest edge of s_m. For any $\epsilon > 0$, we will call σ_m an *ϵ-approximation* to s_m if a homeomorphism, called an *ϵ-homeomorphism*, can be set up between σ_m and s_m such that (1) the distance between any two corresponding points is less than $\epsilon \cdot s_1$ and (2) any tangent m-plane to σ_m has direction cosines (§4) differing by less than ϵ from the corresponding direction cosines of s_m.

LEMMA 2. *For any $\epsilon > 0$, there exists a distance $\rho_3 > 0$ so small that if (Q) is a ϑ-set of diameter less than ρ_3, then σ_m is an ϵ-approximation to s_m.*

The case $m = 1$ is readily treated with the aid of §4(A) and §4, Lemma 2, for in this case $\sigma_1 (\equiv \sigma_m)$ is the image under H^{ik} of a line-segment. Let (Q^0, Q^1) be the end-points of σ_1. Let any point, Q, on σ_1 correspond to the point which divides $Q^1Q^0 (\equiv s_1)$ in the ratio $Q^1Q : Q^1Q^0$. The resulting correspondence will be called a *K_1-correspondence*. It is clearly an ϵ-homeomorphism if s_1 is sufficiently short.

Now let $(Q^0, \cdots, Q^m)$ denote the vertices of σ_m, so ordered that σ_m consists of the[9] $H_1^{i_p}$-cells which join Q^m to the opposite bounding $(m-1)$-cell, σ_{m-1}, of σ_m. Assume Lemma 2 for the value $(m-1)$ and suppose a *K_{m-1}-correspondence*, K_{m-1}, has been defined which is an ϵ-homeomorphism between σ_{m-1} and the $(m-1)$-simplex determined by $(Q^0, \cdots, Q^{m-1})$ provided (Q) is a ϑ-set of sufficiently small diameter. For any point Q on σ_m, let σ_1 denote the $H_1^{i_p}$-cell[9] which joins Q^m to σ_{m-1} and passes through Q. Let Q' be the end-point of σ_1 on σ_{m-1} and let Q'' be the image of Q' under K_{m-1}. To define a *K_m-correspondence*, K_m, let Q correspond to the point which divides the segment Q^mQ'' in the ratio $Q^mQ : Q^mQ'$. We now outline a proof that K_m is an ϵ-homeomorphism if (Q) is a sufficiently small ϑ-set. In the case $j = i_p$, the cell σ_1 is a line-segment, and the tangent m-plane, τ_m, at Q is determined by the line of σ_1 and the tangent $(m-1)$-plane to σ_{m-1} at Q'. If σ_{m-1} is an ϵ'-approximation to s_{m-1}, for ϵ' sufficiently small, then the m-plane τ_m will have[11] direction cosines differing by less than ϵ from those of the m-plane containing s_m. Thus our result follows in this case. If $j = k \neq i_p$, then σ_m is image under H^{jk} of an m-cell σ_m', which falls under the case just treated. Hence, by making (Q) a sufficiently small ϑ-set, we can restrict arbitrarily the variation in the direction cosines on σ_m' and hence on σ_m. We can also make σ_1 an arbitrarily good approximation to the segment Q^mQ'. The tangent m-plane at Q' will then approximate to the m-plane determined by Q^mQ' and the tangent $(m-1)$-plane to σ_{m-1} at Q'. Our result follows readily as in the case $j = i_p$.

LEMMA 3. *If ϵ is sufficiently small in Lemma* 2, *and Q^* is any point of the set (Q) distinct from the vertices of σ_m, then no line through Q^* is tangent to σ_m or meets it in more than one point.*

As a consequence of the definition of θ-set, every line joining Q^* to s_m makes an angle greater than ϑ with s_m. If σ_m is a sufficiently close ϵ-approximation to s_m, every line joining Q^* to σ_m will make an angle greater than $\frac{1}{2}\vartheta$ with s_m, and every chord of σ_m will make an angle less[11] than $\frac{1}{2}\vartheta$ with s_m. Lemma 3 follows immediately.

If ϵ in Lemma 2 satisfies the requirements of Lemma 3, and we make $\rho(\vartheta) > \rho_3$, then (P) will satisfy part (3) of the definition of admissibility [§5(D)]. This follows from the recurrent process used in defining $H_m^{i_1 \cdots i_p}$-cells [§5(B)]. The proof of our theorem is now complete.

7. **Triangulation of L_r.** *Lemma* 1. *For some positive acute angle ϑ it is possible so to triangulate the cells a^i $(i = 1, \cdots, \nu)$ [§3(B)] that the vertices of every r-simplex form a ϑ-set (§6). The triangulation can be made arbitrarily fine.*

Since a^i is a convex polyhedral r-cell in euclidean r-space, this result is easy to establish.

Given a triangulation of a^1, in accordance with Lemma 1, such that every simplex is of diameter less than $\rho(\vartheta)$ [§6, Theorem], let (s^1) denote the subcomplex made up of the closed r-simplexes which meet b^1 [§3(B)]. Under the homeomorphism H^1 [§3(B)] the simplexes (s^1) correspond to a set (σ^1) of H_m^1-cells $(m = 0, \cdots, r)$ [see §5(C)] which satisfy the requirements of the following hypothesis, read for $j = 1$ and $\vartheta_1 = \vartheta$.

HYPOTHESIS. *For some value j in the set $(1, \cdots, \nu - 1)$ there exists a positive angle ϑ_j and a simplicial r-complex (σ^j) on M_r, where (σ^j) is composed of $H_r^{i_1 \cdots i_p}$-cells $(i_1 < i_2 < \cdots < i_p \leqq j)$ and their bounding cells [§5(B)] and has the following properties:*

(1) *Let σ_r be an arbitrary $H_r^{i_1 \cdots i_p}$-cell of (σ^j). Then σ_r has at least one vertex on each of the cells β^k $(k = i_1, \cdots, i_p)$ [§3(A)]. For at least one of these values k, the vertices of σ_r correspond under H^k to a ϑ_j-set of diameter less than $\rho(\vartheta_j)$ [§6, Theorem].*

(2) *For every value $i = 1, \cdots, j$ let $(\sigma^j)^i$ denote the subcomplex of (σ^j) consisting of all the closed r-cells thereof each having at least one vertex on β^i. Then, if the m-cells $(m = 0, \cdots, r)$ of $(\sigma^j)^i$ are replaced by the H_m^i-cells[14] which their vertices determine, the new cells constitute a complex (τ^i) containing all points within distance[9] $\frac{d}{\nu}(\nu - j)$ of γ^i [see §3(B)].*

(3) *Independently of ϑ_j, an arbitrary positive upper bound can be imposed on the diameters of the cells (τ^i).*

We next construct a complex (σ^{j+1}) to satisfy the requirements of the above hypothesis, read with $(j + 1)$ in place of j.

(A) *Making use of hypothesis* (3) [*see also* §4, *Lemma* 2] *we require that if any vertices of an r-cell σ_r of (σ^j) are on any cell β^i of the set (β), then all the vertices of*

[14] The existence of these cells follows from hypothesis (1) and §6, Theorem. See also §5(D).

σ_r shall be on α^i (§3) and they shall correspond under H^i to a set of diameter less than d/ν.

(B) *As a consequence of (A), the complex $(\sigma^j)'$ obtained by deleting all cells of (σ^j) incident with boundary cells thereof on the closure of $M_r - (\sigma^j)$, satisfies all requirements of our hypothesis, provided j is replaced by $(j+1)$ in the expression $\frac{d}{\nu}(\nu - j)$.*

Consider the set of cells $(\sigma^{j+1})^0$ defined as follows, where σ_m denotes any cell of $(\sigma^j)'$. (1) If σ_m has no vertices on β^{j+1}, then σ_m will belong to the set $(\sigma^{j+1})^0$. (2) If σ_m has all its vertices on β^{j+1}, the H_m^{j+1}-cell[14] determined by its vertices will belong to $(\sigma^{j+1})^0$. (3) In any other case, the vertices of σ_m not on β^{j+1} determine an $H_q^{i_1 \cdots i_p}$-cell σ_q $(q < m)$ which falls under case (1), and the remaining vertices determine an H_{m-q-1}^{j+1}-cell σ_{m-q-1} which falls under case (2). The $H_m^{i_1 \cdots i_p,\, j+1}$-cell[14] determined by σ_q and σ_{m-q-1} will belong to $(\sigma^{j+1})^0$.

LEMMA 2. *The cells $(\sigma^{j+1})^0$ form a complex.*

We first obtain an auxiliary result. Let P denote any vertex of $(\sigma^j)'$ on β^{j+1}, and let σ_r be any r-cell of (σ^j) incident with P. As a consequence of our inductive hypothesis, the vertices (P) of σ_r satisfy the hypotheses of the theorem in §6 where $\vartheta = \vartheta_j$. Hence (P) is an admissible set. Let σ_{r-1} be the bounding $(r-1)$-cell of σ_r opposite P. Then σ_{r-1} is an $H_{r-1}^{i_1 \cdots i_p}$-cell, where $(i_1, \cdots, i_p)$ is a subset of $(0, \cdots, j)$ [Hypothesis and §5(A)]. Since (P) is an admissible set, σ_{r-1} and P determine an $H_r^{i_1 \cdots i_p,\, j+1}$-cell, σ_r' [§5(D)]. Let σ_r and its bounding cells be replaced, in the set (σ^j), by σ_r' and its bounding cells. If this is done as σ_r denotes successively each of the r-cells of (σ^j) incident with P, the effect is to replace the cells (σ) of the star of P by a certain set of cells (σ').

(C) *The cells (σ') plus the cells $(\sigma^j) - (\sigma)$ form a complex containing exactly the same points as (σ^j).*

Let π_r denote the set of points on the cells (σ), and let (τ) denote the cells of (σ^j) on the boundary of π_r. We shall write as if[9] each point on α^{j+1} coincided with its image under H^{j+1}. Thus H_1^{j+1}-cells become line-segments and π_r becomes a region of euclidean r-space, covered by (σ) and separated by (τ) from $(\sigma^j) - \pi_r$. The cells (σ') are composed of the line-segments (or H_1^{j+1}-cells) joining P to (τ) [§5(B)]. Since the line-segment joining P to any point of π_r can be continued to meet (τ), all the points of π_r are on (σ'). If Q is on (τ), then a neighborhood of Q on the line-segment PQ is on[15] π_r. Hence (1) every segment joining P to (τ) is on the closure of π_r and (2) the segments joining a neighborhood of Q on (τ) to the point P constitute a cone containing a neighborhood on M_r of every interior point of PQ. Since, by (1), every point of this cone is on the closure of π_r, it follows that every interior point of PQ is on π_r.

---

[15] Otherwise PQ approaches some $(r-1)$-cell σ_{r-1} of (τ) from the outer side relative to the cells (σ); and PQ, by definition, is on one of the cells (σ'). But, in the paragraph preceding (C), the inner side of σ_{r-1} relative to σ_r must be the same as its inner side relative to σ_r'. This follows as a corollary to the lemmas in §6.

This completes a proof that the set of all points on (σ') coincides with the set of points on (σ). Furthermore, since no segment from P can meet (τ) in more than one point, the cells (σ') are distinct. Result (C) now follows directly.

It is possible[16] to obtain $(\sigma^{j+1})^0$ from (σ^j) by modifying successively the stars on (σ^j) of all the vertices of $(\sigma^j)'$ which lie on β^{j+1} precisely as the star of P was modified, and then dropping all cells with limit points on the closure of $M_r - (\sigma^j)$. The proof of result (C) applies to show that after each of these successive modifications we still have a complex.

(D) *The complex* $(\sigma^{j+1})^0$ *satisfies the requirements of the hypothesis if* j *is replaced by* $(j+1)$ *in the expression* $\frac{d}{\nu}(\nu - j)$ *and in the conditions* $(i_1 < i_2 < \cdots < i_p \leqq j)$.

This follows from result (B) and the definition of $(\sigma^{j+1})^0$.

It remains to secure hypothesis (2) for the additional value $i = j + 1$ and, if necessary, to replace ϑ_j by a smaller angle ϑ_{j+1}. Let (s^{j+1}) be a triangulation of α^{j+1} in accordance with Lemma 1 above, each simplex being of diameter less than d/ν [§3(B)]. Let $(\sigma^{j+1})''$ denote the complex on α^{j+1} consisting of the closed H_r^{j+1}-cells each of which corresponds to one of the simplexes (s^{j+1}) and contains a point of $M_r - (\sigma^{j+1})^0$ within distance d of γ^{j+1}. Since the cells $(\sigma^{j+1})''$ are entirely within distance $\left(1 + \frac{1}{\nu}\right) d$ of γ^{j+1}, the only cells of $(\sigma^{j+1})^0$ they can meet are the H_m^{j+1}-cells. Hence, by a finite number of elementary subdivisions,[17] we can subdivide $(\sigma^{j+1})^0$ and $(\sigma^{j+1})''$, using only H_m^{j+1}-cells, into complexes with just a sub-complex in common. The totality of cells in the two subdivisions then satisfies the hypothesis, read for $(j+1)$, save perhaps for the requirements involving ϑ_j. We can satisfy these also, for some angle ϑ_{j+1}, by slightly displacing certain vertices involved in the elementary subdivisions and imposing a new upper bound on all the cells we have used. We then have a complex (σ^{j+1}) which satisfies the entire hypothesis, read with $(j+1)$ in place of j.

Starting with (σ^1) *and making* $(\nu - 1)$ *successive applications of the above process, we obtain a triangulation* (σ^ν) *of* M_r. *The extension of such a triangulation to the other elements of* L_r *presents no new difficulties.*

LEHIGH UNIVERSITY.

---

[16] See the definition of $H_m^{i_1 \cdots i_p}$-cells in §5, where the vertices are introduced one at a time.

[17] Lefschetz, Topology (1930) pp. 65ff.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

DISCRETE GROUPS GENERATED BY REFLECTIONS

BY H. S. M. COXETER

(Received June 16, 1933)

CONTENTS

1–3. Polytopes with restricted dihedral angles 590
4–7. Fundamental regions 594
8. The abstract definition 599
9. Complete enumeration of the irreducible groups 601
10–13. Continued products of the generators 602
14, 15. The central inversion 606
16. The total number of reflections 610
17. Generation by two operations 612
18. Elliptic space 617
Table of irreducible finite groups generated by reflections 618
Graphical representation 619
Index 620
References 620

Introduction

It is known that, in ordinary space, the only finite groups generated by reflections are

$[k]$ $(k \geqq 1)$, of order $2k$, with abstract definition

$$R_1^2 = R_2^2 = (R_1R_2)^k = 1,$$

and

$[k_1, k_2]$ $(k_1 \geqq 2,\ k_2 \geqq 2,\ 1/k_1 + 1/k_1 > \frac{1}{2})$, of order $\dfrac{4}{1/k_1 + 1/k_2 - \frac{1}{2}}$, with abstract definition

$$R_1^2 = R_2^2 = R_3^2 = (R_1R_2)^{k_1} = (R_1R_3)^2 = (R_2R_3)^{k_2} = 1.$$

$[1]$ is the group of order 2 generated by a single reflection. Since this is the symmetry group of the one-dimensional polytope[1] $\{\}$, we write

$$[1] = [\,].$$

$[k](k \geqq 3)$ and $[k_1, k_2]$ $(k_1 = 3,\ k_2 = 3, 4, 5)$ are the symmetry groups of the ordinary regular polygons $\{k\}$ and polyhedra $\{k_1, k_2\}$. The rest of the groups can be written as direct products, thus:

$$[2] = [\,] \times [\,],$$
$$[2, 2] = [\,] \times [\,] \times [\,],$$
$$[2, k] = [\,] \times [k].$$

[1] Coxeter 1, 344.

These groups can be made vividly comprehensible by using actual mirrors for the generating reflections. It is found that a candle makes an excellent object to reflect. By hinging two vertical mirrors at an angle π/k ($k = 2, 3, 4, \ldots$), we easily see $2k$ candle flames, in accordance with the group $[k]$. To illustrate the groups $[k_1, k_2]$, we hold a third mirror in the appropriate positions.

In the present paper, we generalize the above results to space of any number of dimensions,[2] and to groups which, though infinite, are still free from infinitesimal operations. (One such infinite group is called $\boxed{3^3}$, and may be illustrated by means of three vertical mirrors, erected on the sides of an equilateral triangle so as to form a prism, open at the top. A candle placed within this prism gives rise to an unlimited number of images.)

In order to enumerate these groups, we observe that the fundamental region[3] must be a polytope whose dihedral angles, being submultiples of π, are never obtuse. It happens that such polytopes are of a particularly simple form.

Perhaps it is worth while to point out what contact this investigation has with crystallography. Each of the 32 crystal systems corresponds to a finite group of orthogonal transformations. Of those groups, the eleven which Schoenflies[4] calls *Holoedrie* and *Hemimorphe Hemiedrie* are generated by reflections; in fact,

$$
\begin{aligned}
C_1^h &= [\,], \\
C_2^v &= [\,] \times [\,], \\
V^h &= [\,] \times [\,] \times [\,], \\
C_k^v &= [k] \quad (k = 3, 4, 6), \\
D_k^h &= [\,] \times [k], \\
T^d &= [3, 3], \\
O^h &= [3, 4],
\end{aligned}
$$

Of the 230 *space groups*, the following seven are generated by reflections:

$$
\begin{aligned}
\mathfrak{B}_h^1 &= [\infty] \times [\infty] \times [\infty], \\
\mathfrak{D}_{6,h}^1 &= [\infty] \times [3, 6], \\
\mathfrak{D}_{4,h}^1 &= [\infty] \times [4, 4], \\
\mathfrak{D}_{3,h}^1 &= [\infty] \times \boxed{3^3},[5] \\
\mathfrak{T}_h^3 &= \boxed{3^4}, \\
\mathfrak{O}_h^5 &= \begin{bmatrix} 4 \\ 3 \\ 3 \end{bmatrix},[6] \\
\mathfrak{O}_h^1 &= [4, 3, 4].[7]
\end{aligned}
$$

[2] For the case of four dimensions, see Goursat **1**, 80–93.
[3] Bieberbach **1**, 312.
[4] Schoenflies **1**.
[5] Coxeter **2**, 147.
[6] *Ibid.*, 148.
[7] *Ibid.*, 150.

([∞] is the group generated by reflections in two points of a line. Its abstract definition is simply

$$R_1^2 = R_2^2 = 1.)$$

THEOREM 1. *In spherical space, every polytope free from obtuse dihedral angles is a simplex.*

LEMMA 1.1. *Every spherical polygon free from obtuse angles is a triangle.*

Since the angle-sum of an n-gon is greater than $(n - 2)\pi$, at least one angle must be greater than $(n - 2)\pi/n$.

LEMMA 1.2. *If two angles of a spherical triangle are non-obtuse, then the third angle is not less than its opposite side.*

By the well-known formula, if A and B are non-obtuse,

$$\cos C = \sin A \sin B \cos c - \cos A \cos B \leqq \cos c.$$

DEFINITION. A *simplex-polycorypha*[8] is a polytope in which the number of edges at any vertex is equal to the number of dimensions of the space. *E.g.*, the measure polytope is a simplex-polycorypha.

LEMMA 1.3. *If a simplex-polycorypha is free from obtuse dihedral angles, so is any bounding figure.*

Consider the section of the simplex-polycorypha Π_m by a 3-space perpendicular to any element Π_{m-3}, meeting that element at an internal point O. The three Π_{m-1}'s which meet at the Π_{m-3} will give, in the section, a trihedral angle with vertex O; and the edges of this trihedral angle will be sections of the three Π_{m-2}'s which separate the Π_{m-1}'s in pairs. Further, on account of the orthogonality, the dihedral angles of the trihedral angle are dihedral angles of Π_m, while its face-angles are dihedral angles of the three Π_{m-1}'s. On a sphere with center O, the trihedral angle cuts out a spherical triangle whose angles A, B, C are dihedral angles of Π_m, and so non-obtuse, by hypothesis. Lemma 1.2 gives

$$c \leqq C \leqq \tfrac{1}{2}\pi.$$

But c can be any dihedral angle of any bounding figure. Thus Lemma 1.3 is proved.

(If $m = 3$, there is no need to take a section. If $m < 3$, the Lemma is meaningless.)

LEMMA 1.4. *In three or more dimensions, every simplex-polycorypha bounded entirely by simplexes is itself a simplex.*

In an m-dimensional simplex-polycorypha, $m - 1$ plane faces pass through each edge. If all the bounding figures are simplexes, these plane faces are triangles. Let $A_1, A_2, \ldots, A_m$ be the m vertices that are joined to the vertex A_0 by edges. Then the $m - 1$ triangles through A_0A_1 must be

$$A_0A_1A_2, \ldots, A_0A_1A_m.$$

---

[8] Sommerville 1. For the *polytopes* considered here, the definition given in Coxeter 1, 331 (§1.1) is appropriate. Sommerville calls these *simple convex* polytopes.

By hypothesis, just m edges pass through A_1. Therefore A_1A_0, A_1A_2, ..., A_1A_m are the *only* edges through A_1. Similarly for any other A_i. Thus the polycorypha has only $m + 1$ vertices.

Theorem 1 can now be proved by induction. By Lemma 1.1, it is true in two dimensions. Let us then assume it true in $m - 1$ dimensions. Consider an m-dimensional polytope free from obtuse dihedral angles. A small sphere[9] drawn round any vertex cuts out an $(m - 1)$-dimensional spherical polytope whose dihedral angles all occur among those of the m-dimensional polytope. By the inductive assumption, this $(m - 1)$-dimensional polytope is a simplex; *i.e.*, the m-dimensional polytope is a simplex-polycorypha. By Lemma 1.3 and the inductive assumption, every bounding figure is a simplex. The theorem follows by Lemma 1.4.

COROLLARY. *In Euclidean space, every acute-angled polytope is a simplex.*

If possible, consider an acute-angled polytope that is not a simplex. By projection on to a spherical space of sufficiently large radius, we can obtain a spherical polytope whose dihedral angles differ by as little as we please from those of the given polytope. If we make the radius so large that the angles remain acute, Theorem 1 is contradicted.

DEFINITION. A *prism*[10] is the topological product of a number of polytopes lying in absolutely perpendicular spaces. Regarding a polytope as a closed set of points (*viz.* all the points within it and on its boundary), we can define the prism $[\Pi_p, \Pi_q]$ as follows.

In $p + q$ dimensions, suppose Π_p to be fixed, while Π_q is constrained to lie in an absolutely perpendicular space and to have a constant orientation. Suppose further that the common point of the spaces of Π_p and Π_q belongs to both polytopes, and is a *definite* point of Π_q (*e.g.* a vertex). Then the prism is the totality of points that belong to the possible positions of Π_q.

Topological products being commutative and associative, a prism may have any number of constituents Π_p, Π_q, ..., and the order of their arrangement is immaterial. It is convenient to admit the trivial case when there is only one constituent.

A *simplicial prism* is a "rectangular product" of (one or more) simplexes. We use the symbol Σ_m for a general simplex in m dimensions. Thus Σ_0 denotes a point, and Σ_1 a straight segment (of unspecified length).

THEOREM 2. *In Euclidean space, every polytope free from obtuse dihedral angles is a simplicial prism.*

LEMMA 2.1. *Every polygon free from obtuse angles is either a triangle Σ_2 or a rectangle $[\Sigma_1, \Sigma_1]$.*

This follows from the fact that the angle-sum of an n-gon is $(n - 2)\pi$.

LEMMA 2.2. *Every polytope free from obtuse dihedral angles is a simplex-polycorypha.*

---

[9] We shall use the word *sphere* for the analogue in any number of dimensions.

[10] Called *prismotope* by Schoute **1**. *Cf.* Coxeter **1**, 351.

A small sphere drawn round any vertex cuts out a spherical polytope whose dihedral angles all occur among those of the whole polytope. By Theorem 1, this spherical polytope is a simplex.

LEMMA 2.3. *A prism whose constituents are simplex-polycoryphas is itself a simplex-polycorypha.*

Since every vertex of $[\Pi_p, \Pi_q]$ is a vertex of one Π_p and of one Π_q, the edges which meet there are the covertical edges of Π_p and of Π_q. If, in particular, Π_p and Π_q have respectively p and q covertical edges, then $[\Pi_p, \Pi_q]$ has $p + q$; *i.e.* the prism is a simplex-polycorypha. This result, having been proved for two constituents, at once extends to the case of any number.

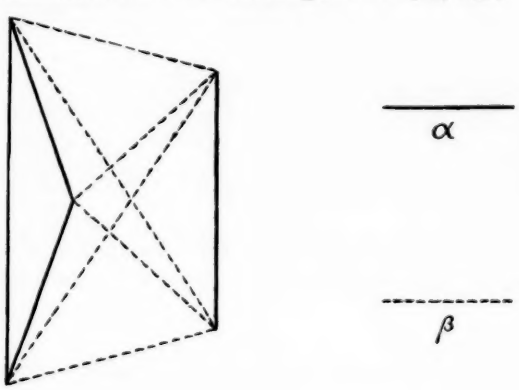
LEMMA 2.4. *Every simplicial prism can be characterized by a "vertex diagram" consisting of a simplex with edges of two types.*

Consider again the construction for $[\Pi_p, \Pi_q]$. When Π_q moves so that each point of it describes an edge of Π_p, its edges describe rectangles. Thus, if AB and AD are edges of Π_p and Π_q respectively, the rectangle $ABCD$ is an element of $[\Pi_p, \Pi_q]$.

In the simplicial prism $[\Sigma_p, \Sigma_q, \ldots]$, the edges that meet at any vertex fall into sets of p, q, etc., such that the first set belong to Σ_p, the second set to Σ_q, and so on. Two edges occurring in any one set are sides of a triangle belonging to the corresponding simplex. Two edges occurring in separate sets are sides of a rectangle.

We can now represent the $p + q + \ldots$ covertical edges as vertices of a topological simplex, putting the mark "α" against those edges of the simplex which correspond to triangles, and "β" against those which correspond to rectangles. The vertices of this *vertex diagram* fall into sets of $p, q, \ldots$, such that every pair occurring in the same set are joined by α-edges, while every pair occurring in different sets are joined by β-edges. (In the special case of $[\alpha_p, \alpha_q, \ldots]$,[11] the vertex diagram can be identified with the *vertex figure*, and then the α-edges and β-edges are α_1's and β_1's;[12] hence the notation.)

Since the form of the vertex diagram depends only on the numbers $p, q, \ldots$, we shall not obtain a different diagram by beginning with a different vertex. (The following illustration is the vertex diagram for $[\Sigma_3, \Sigma_2]$.)



[11] α_p means *a regular* Σ_p *of unit edge.*

[12] Coxeter **1**, 345.

LEMMA 2.5. *Every trihedral*[13] *polyhedron bounded by triangles* Σ_2 *and rectangles* $[\Sigma_1, \Sigma_1]$ *is either a tetrahedron* Σ_3 *or a triangular prism* $[\Sigma_2, \Sigma_1]$ *or a rectangular solid* $[\Sigma_1, \Sigma_1, \Sigma_1]$.

If a trihedral polyhedron is bounded by f_3 triangles and f_4 quadrangles, it must have $\frac{1}{2}(3f_3 + 4f_4)$ edges and $\frac{1}{3}(3f_3 + 4f_4)$ vertices. Hence $\frac{1}{2}f_3$, $\frac{1}{3}f_4$ must be (non-negative) integers, and, by Euler's Theorem,

$$\tfrac{1}{2}f_3 + \tfrac{1}{3}f_4 = 2.$$

The solutions $2 + 0, 1 + 1, 0 + 2$ give the three possible kinds of polyhedron.

LEMMA 2.6. *If a simplex-polycorypha in more than three dimensions is bounded solely by simplicial prisms, then the types of prisms occurring at any vertex are the same as at any other vertex.*

If a simplex-polycorypha in m dimensions is bounded solely by simplicial prisms, we can construct a vertex diagram at any vertex, and we have to show that this will be the same for all vertices. Any edge AB of the simplex-polycorypha will be represented by a vertex B' in the vertex diagram at A and by a vertex A' in the vertex diagram at B. The edge AB belongs to $m - 1$ bounding prisms; in the vertex diagrams considered, these are represented by the $m - 1$ bounding simplexes that meet at B' or A' respectively. Hence the two vertex diagrams are simplexes, alike in $m - 1$ of their m bounding figures. If $m > 3$, this is sufficient to make them identical.

Since any vertex of the polycorypha can be reached from any other by a chain of edges, repeated application of this result proves the lemma.

LEMMA 2.7. *In three or more dimensions, every simplex-polycorypha bounded solely by simplicial prisms is itself a simplicial prism.*

The case of three dimensions is covered by Lemma 2.5, so we may suppose the number of dimensions to exceed three. By Lemma 2.6, the polycorypha is characterized by a vertex diagram, *viz.* a simplex with α-edges and β-edges. Let P be any vertex of this vertex diagram. The remaining vertices form a simplex which, being the vertex diagram of one of the bounding prisms of the polycorypha, has sets of vertices joined among themselves by α-edges and joined to one another by β-edges.

We have to prove that the vertices of the whole simplex fall into sets in the same manner. This is certainly true if all the edges through P are of type β. On the other hand, if P is joined to one of the other vertices by an α-edge, it must be joined by α-edges to all the other vertices of the same set, since otherwise there would be a triangle of sides α, α, β, which is impossible by Lemma 2.5. For the same reason, P cannot be joined by α-edges to vertices of two distinct sets. Thus the vertices of the whole simplex are distributed in the desired manner, the vertex P either forming a new set by itself or becoming attached to one of the old sets.

LEMMA 2.8. *Every simplex-polycorypha free from obtuse dihedral angles is a simplicial prism.*

By Lemma 2.1, this is true in two dimensions. Let us assume it true in $m - 1$ dimensions, and use induction. By Lemma 1.3 and the inductive assumption,

[13] *I.e.*, a polyhedron with trihedral vertices.

the m-dimensional polycorypha is bounded by simplicial prisms. Hence, by Lemma 2.7, it is itself a simplicial prism.

Theorem 2 follows from Lemmas 2.2 and 2.8.

THEOREM 3.[14] *If all the dihedral angles of an m-dimensional polytope are less than or equal to θ $(0 < \theta < \pi)$, then the number of bounding figures is less than some number depending only on m and θ.*

Take first the case when $m = 3$. In the spherical image by parallel normals ("Gaussian" image for surfaces of continuous curvature), to each face of the polyhedron corresponds to a point on the unit sphere, to an edge where the dihedral angle is θ_1 corresponds a great-circular arc of length $\pi - \theta_1$, and to an n-hedral vertex corresponds a spherical n-gon. By hypothesis, the spherical distances between the points on the sphere are greater than or equal to $\pi - \theta$. Consequently, circles of spherical radius $\frac{1}{2}(\pi - \theta)$, described round the points, will not overlap. If N is the number of faces of the polyhedron, comparison of areas gives

$$2\pi(1 - \sin \tfrac{1}{2}\theta)N < 4\pi,$$

$$N < 2/(1 - \sin \tfrac{1}{2}\theta).$$

When $\theta = \frac{1}{2}\pi$, this gives the "best possible" result $N \leqq 6$. (The actual polyhedra are those considered in Lemma 2.5.)

The analogous result for general m is easily seen to be

$$N < 2 \int_0^{\frac{1}{2}\pi} \sin^{m-2}\varphi \, d\varphi \Big/ \int_0^{\frac{1}{2}(\pi-\theta)} \sin^{m-2}\varphi \, d\varphi.$$

E.g., for $m = 4$, $N < 2\pi/(\pi - \theta - \sin\theta)$.

DEFINITIONS. A group of congruent transformations (of Euclidean space into itself) is said to be *discrete* if the totality of transforms of a point never has a limit point. *E.g.*, the group generated by a single rotation, through an angle incommensurable with π, is not discrete; but every finite group is *a fortiori* discrete. Bieberbach[15] has proved that a group is discrete if and only if it is free from infinitesimal operations.

Consider a point of such general position that it has distinct transforms under all the different operations of the group. Sufficiently small neighborhoods of these transforms can be taken so as to be equivalent under the group, and not overlap. By gradually increasing these neighborhoods, we eventually obtain a set of congruent[16] regions which together fill up the whole space. Any one of these regions is, by definition, a *fundamental region*; and the group can be generated by the operations which transform this particular region into its neighbors.

---

[14] I am indebted to Prof. G. Pólya for this extension of Theorem 2. It will not be applied in the sequel, but is inserted for its intrinsic interest.

[15] Bieberbach **1**, 313–314.

[16] In the wide sense; *i.e.*, directly congruent, or enantiomorphous.

A point of general position on the boundary of the fundamental region is also on the boundary of a neighboring region, and may or may not be transformed into itself by the relevant generator. If not, a sufficiently small neighborhood of this point, in the neighboring region, will be equivalent to a certain portion of the original region. The fundamental region can then be modified by subtracting this portion and adding the aforesaid neighborhood. Thus the only case in which the fundamental region is uniquely determined is when every point of the boundary is transformed into itself. It is easily seen that this can only happen when all the generators are reflections.

A group of congruent transformations is said to be *reducible*[17] if it can be regarded as the direct product of groups of congruent transformations in two absolutely perpendicular complementary subspaces $\bar{\omega}$, $\bar{\omega}'$. Every operation is then of the form QQ' $(=Q'Q)$, where Q transforms $\bar{\omega}$ into itself and Q' transforms $\bar{\omega}'$ into itself. The groups in the subspaces are called *components*, and may themselves be reducible; but we can eventually analyse the original group into a number of *irreducible* components.

A group is said to be *trivially reducible* if it has identity for a component. In this case the complementary component operates in a subspace $\bar{\omega}$ which is transformed into itself by every operation of the group, and we regard the fundamental region as lying in $\bar{\omega}$ instead of in the whole space. In other words, we consider the section of the true fundamental region by $\bar{\omega}$. Thus we regard a trivially reducible group and its non-trivial component as having the same fundamental region.

If a group leaves one point invariant, it transforms into itself the unit sphere around this point. In this case we regard the fundamental region as lying on the sphere. In other words, we replace the angular fundamental region by its spherical section. Thus we confuse a group in spherical space with the corresponding Euclidean group which leaves the center invariant. (A group which leaves more than one point invariant is always trivially reducible.)

Consider, for example, the group generated by reflections in two perpendicular planes in ordinary space. This is trivially reducible, its non-trivial component is generated by reflections in two perpendicular lines of a plane, and its fundamental region is an arc of length $\frac{1}{2}\pi$.

We use the word *prime* for a space of one fewer than the current number of dimensions, and *secundum* for the next lower space (*i.e.* the intersection of two primes).

The Fricke-Klein construction for a fundamental region.[18] Given a point P which is not invariant under any operation, and its transforms $P_1, P_2, \ldots$, we can construct a fundamental region by drawing the perpendicularly bisecting

---

[17] More strictly, "completely reducible." When the group is finite, our omission of the word "completely" is justified by Burnside **1.**

[18] Fricke-Klein **1,** 108, 216.

primes of PP_1, PP_2, ... A finite[19] number of these primes cut off a fundamental region around P, and the rest of them are irrelevant.

ABBREVIATION. For "group generated by reflections" we shall write "g.g.r.".

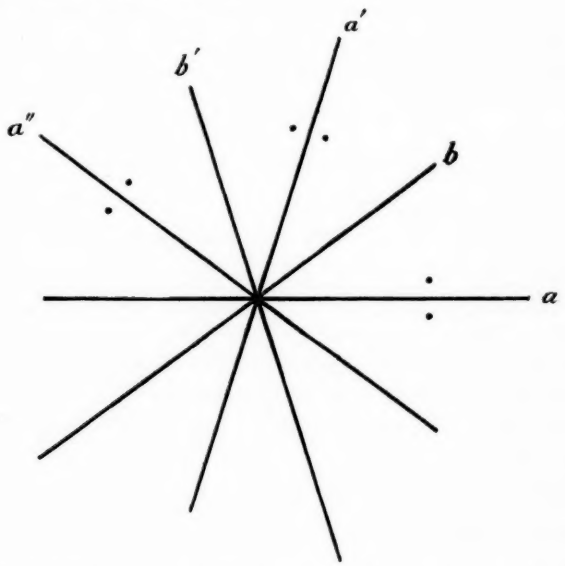
THEOREM 4. *The general finite g.g.r. has for fundamental region a spherical simplex, whose vertices fall into sets which belong to the fundamental regions of the irreducible components, vertices in different sets being joined by edges of length* $\frac{1}{2}\pi$.

LEMMA 4.1.[20] *Every finite group of congruent transformations leaves at least one point invariant.*

Consider the set of transforms of a point of general position. Every operation of the group permutes these points among themselves, and leaves their centroid invariant.

LEMMA 4.2. *If a discrete group is generated by reflections in certain primes, then these primes and their transforms are so distributed that, whenever n of the primes meet in a secundum, they are inclined to one another at angle* π/n *in cyclic succession.*

Let a, b be any two of the primes that are not parallel. Then the group must contain reflections in new primes a', b', a'', b'', . . . which are the reflected images of a, b, a', b', . . . in b, a', b', a'', . . . respectively. These are inclined to a at all multiples of the angle $(a\,b)$.



If this angle were incommensurable with π, the sequence of primes would be infinite, a point of general position would have an infinite number of transforms lying on a finite circle, and the group would not be discrete. Hence we can write

$$(a\,b) = d\,\pi/n,$$

[19] Cartan **1**.

[20] Bieberbach **1**, 327 (XI). Coxeter **2**, 182 (§20.2) is an immediate corollary of this lemma.

where n and d are co-prime integers. If integers ν and δ are chosen so that $\nu d - \delta n = 1$, we shall have

$$\nu (a\ b) = \delta \pi + \pi/n.$$

Thus there is a prime inclined to a at angle π/n, and the primes can be re-ordered so as to make $d = 1$.

LEMMA 4.3. *If a finite group is reducible, there exist two absolutely perpendicular complementary subspaces $\bar{\omega}$, $\bar{\omega}'$, such that every operation is of the form QQ', where Q leaves invariant every point of $\bar{\omega}'$ and Q' leaves invariant every point of $\bar{\omega}$.*

In the notation of matrices,

$$\begin{pmatrix} A & 0 \\ 0 & A' \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & A' \end{pmatrix}.$$

LEMMA 4.4. *If a reducible finite group is generated by reflections, its components are likewise generated by reflections.*

Consider any reflection which belongs to a reducible finite group. By Lemma 4.3, it transforms $\bar{\omega}$ and $\bar{\omega}'$ into themselves. Therefore $\bar{\omega}$ (or $\bar{\omega}'$) must lie in the reflecting prime, while $\bar{\omega}'$ (or $\bar{\omega}$) is perpendicular to it. Hence Q (or Q') is identity, while Q' (or Q) is a reflection.

LEMMA 4.5. *If a finite g.g.r. is not trivially reducible, it leaves invariant just one point of the Euclidean space, and its fundamental region is one of the parts into which the unit sphere around this point is divided by the reflecting primes and their transforms.*

A finite g.g.r., operating in Euclidean space, leaves at least one point invariant, by Lemma 4.1. If it leaves more than one point invariant, it must leave all the points of a subspace of one or more dimensions; this makes it trivially reducible. The group can be regarded as operating in a spherical space, and its fundamental region is then given by the Fricke-Klein construction.

LEMMA 4.6. *If a finite g.g.r. is not trivially reducible, it is generated by reflections in the bounding primes of its fundamental region.*

These are the operations which transform the fundamental region into the neighboring regions.

LEMMA 4.7. *The fundamental region of a finite g.g.r. is a spherical simplex.*

If the fundamental region has a pair of antipodal vertices, all its bounding primes must pass through the join of these vertices, since any other prime would divide the region into two parts. There are thus two invariant points, the group is trivially reducible, and we construct the fundamental region in a subspace.

If the fundamental region has no vertices (*e.g.* a hemisphere), suppose it has a k-dimensional element but no $(k-1)$-dimensional element. Then all the bounding primes must pass through this k-dimensional element, since any other prime would give a $(k-1)$-dimensional element by intersection.

Thus the fundamental region, which we have agreed to derive from the non-trivial component, has vertices, but no antipodal vertices. This makes it a

polytope. By Lemma 4.2, all the dihedral angles are submultiples of π; therefore none of them can be obtuse. By Theorem 1, the polytope is a simplex.

The rest of Theorem 4 follows by means of Lemma 4.4. The sets of vertices lie in absolutely perpendicular spaces.[21]

THEOREM 5. *If an infinite discrete g.g.r. has a finite fundamental region, this fundamental region is a simplicial prism whose constituents are the fundamental regions of the irreducible components.*

LEMMA 5.1. *If a reducible discrete group is generated by reflections, its components are likewise generated by reflections.*

We proceed as in Lemma 4.4, bearing in mind the fact that the subspaces $\tilde{\omega}$, $\tilde{\omega}'$ may be at infinity.

LEMMA 5.2. *If an infinite discrete g.g.r. is not trivially reducible, its fundamental region is one of the parts into which space is divided by the reflecting primes and their transforms. It is generated by reflections in the bounding primes of the fundamental region.*

Cf. Lemmas 4.5, 4.6.

The fundamental region of an infinite discrete g.g.r. may be finite or infinite. If finite, Lemma 4.2 shows that it is a Euclidean polytope whose dihedral angles are submultiples of π. Hence, by Theorem 2, it is a simplicial prism.

The bounding figures of the simplicial prism

$$[\Sigma_{m_1}, \Sigma_{m_2}, \Sigma_{m_3}, \ldots]$$

fall into sets of the form

$$[\Sigma^{(i)}_{m_1-1}, \Sigma_{m_2}, \Sigma_{m_3}, \cdots], \qquad [\Sigma_{m_1}, \Sigma^{(j)}_{m_2-1}, \Sigma_{m_3}, \cdots], \text{ etc.}$$

Two bounding figures belonging to different sets lie in perpendicular primes, so that the reflections in them are permutable. The group generated by reflections in the bounding primes of the prism is thus the direct product of groups generated by reflections in the bounding primes of Σ_{m_1}, Σ_{m_2}, etc. In this manner we analyse the group into its irreducible components.

COROLLARY. *All the components are then infinite.*

THEOREM 6. *Every irreducible discrete g.g.r. has a simplicial fundamental region; and the general discrete g.g.r. is simply isomorphic with a direct product of such groups.*

If the group is finite, the fundamental region is a simplex (whether the group is irreducible or not). If the group is infinite while the fundamental region is finite, the fundamental region is a simplicial prism; but the group is reducible if this prism has more than one constituent. Bieberbach[22] has proved that groups

---

[21] Cf. Coxeter **2**, 146 (§16.4). The fundamental region of $[m] \times [n]$ is described in Goursat **1**, 82.

[22] Bieberbach **1**, 327.

with infinite fundamental regions are always reducible. By Lemma 5.1, the irreducible components are generated by reflections.

The words *"simply isomorphic with"* are inserted in order to admit trivially reducible groups.

THEOREM 7. *If a discrete g.g.r. has both finite and infinite components, its fundamental region can be regarded as a generalized pyramid which joins a simplicial prism to a simplex situated at infinity in a space absolutely perpendicular to the space of the prism.*

Such a group is the direct product of a finite group whose fundamental region is a spherical simplex, and an infinite group whose fundamental region is a simplicial prism. Let us imagine the spherical simplex to be drawn (with its proper angles) on a sphere of infinite radius, lying in a Euclidean $(q+1)$-space absolutely perpendicular to the p-space of the prism. Consider the infinite $(p+q+1)$-dimensional pyramidal region obtained by joining every point of the prism to every point of the simplex at infinity. This region has bounding primes of two types:

(1) joining a bounding $(p-1)$-space of the prism to the q-space (at infinity) of the simplex;

(2) joining a bounding $(q-1)$-space of the simplex to the p-space of the prism.

Since the p-space is absolutely perpendicular to the q-space at infinity, every prime of the first type is perpendicular to every prime of the second. Reflections in the primes of the first type generate the infinite component of our group, while reflections in primes of the second type generate the finite component. Hence reflections in all the primes must generate the direct product, as required.

For example, the fundamental region of

$$[k] \times [\infty]$$

is an infinite "wedge" (in ordinary space), bounded by four planes: two parallel; and two, perpendicular to these, mutually inclined at π/k.

THEOREM 8. *Every discrete g.g.r. has an abstract definition of the form*

$$R_i^2 = (R_i R_j)^{k_{ij}} = 1. \tag{8.1}$$

Since the direct product of several groups having such abstract definitions is another group of the same kind, it will be sufficient, by Theorem 6, to consider the cases when the fundamental region is a spherical or Euclidean simplex. Let R_i denote the reflection in the i^{th} bounding prime of this simplex, and π/k_{ij} the angle between the i^{th} and j^{th} bounding primes. Then the relations (8.1) are certainly satisfied. Clearly, also, the R's suffice to generate the whole group. It remains to be proved that every relation satisfied by the R's is a consequence of (8.1).

Let[23]

$$R_a R_b R_c \ldots R_z = 1 \tag{8.2}$$

[23] This proof is the "obvious extension of an argument used by Burnside" to which reference was made in Coxeter **2**, 146.

be any relation satisfied by the R's. Consider a point, moving continuously in the spherical or Euclidean space in which the given group operates, beginning within a particular fundamental region, proceeding through the a-face into a neighboring region, then through the b-face of this region into a third (which may be the first over again), ... and finally through the z-face of some region into another region (which we shall identify with the first).

The first region is transformed into the second by the operation R_a, into the third by

$$R_a R_b R_a^{-1} . R_a = R_a R_b,$$

into the fourth by

$$(R_a R_b) R_c (R_a R_b)^{-1} \cdot R_a R_b = R_a R_b R_c,$$

... and into the last by $R_a R_b R_c \ldots R_z$. By hypothesis, this operation is identity. Hence the last region is the same as the first, and we may regard the moving point as describing a *closed* path.

We can immediately reduce the general sequence

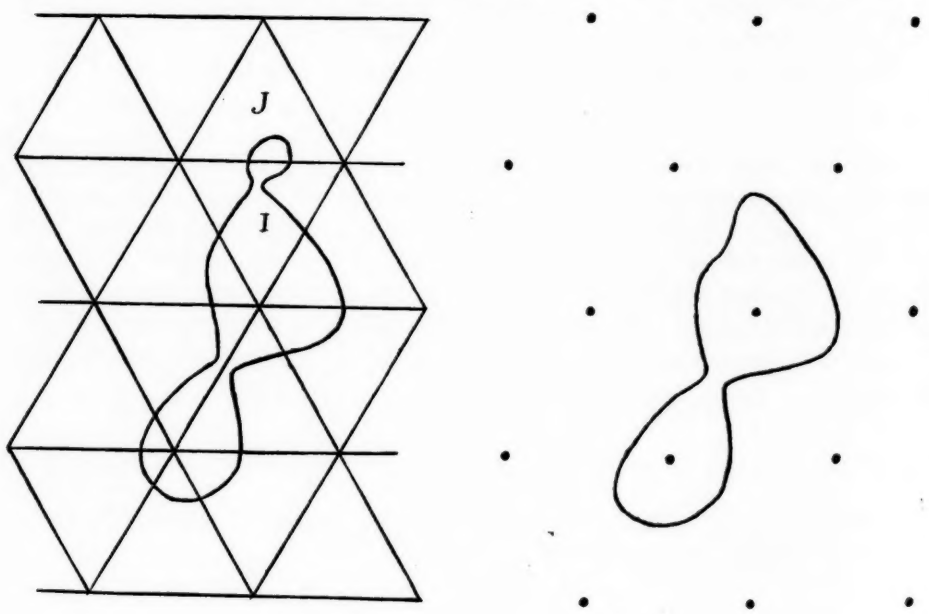
$$R_a, R_b, \ldots, R_z$$

to one in which no two consecutive R's are the same. For, since

$$R_i^2 = 1,$$

$$R_a \ldots R_h R_i R_j R_k \ldots R_z = R_a \ldots R_h R_k \ldots R_z \quad \text{if} \quad i = j.$$

In terms of the path, the moving point passes through the i-face of some region I into J (say), and then back through the same interface into I. This path can be decomposed into one which misses J altogether (unless it enters on a different occasion) along with a loop passing twice through the interface. The loop itself is a path which corresponds, in the manner described, to the relation $R_i^2 = 1$.



We shall now reduce the net of fundamental regions to a "skeleton," by discarding the primes but retaining the secunda. The possible paths, of the kind considered, constitute the *Poincaré group* of the space residual to these secunda. Since the whole space is simply connected, any such path can be decomposed into loops around single secunda. This decomposition of the path corresponds to a factorization of $R_a R_b \dots R_z$. By reinserting the primes, we see that each factor is of the form

$$(R_p \dots R_q)\,(R_i R_j)^{kij}\,(R_p \dots R_q)^{-1}.$$

Hence, finally, (8.2) is an algebraic consequence of

$$(R_i R_j)^{kij} = 1.$$

THEOREM 9. *The following list*[24] *comprises all irreducible discrete groups generated by reflections.*

Finite groups: $[3^n]$ $(n \geqq 0)$, $[3^n, 4]$ $(n \geqq 0)$, $\begin{bmatrix} 3^n \\ 3 \\ 3 \end{bmatrix}$ $(n \geqq 1)$,[25] $[k]$ $(k \geqq 5)$,

$$[3,5], \quad [3,4,3], \quad [3,3,5], \quad \begin{bmatrix} 3,3 \\ 3,3 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 3,3,3 \\ 3,3 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 3,3,3,3 \\ 3,3 \\ 3 \end{bmatrix}.$$

Infinite groups: $\boxed{3^n}$ $(n \geqq 3)$, $[4,3^n,4]$ $(n \geqq 0)$, $\begin{bmatrix} 3^n,4 \\ 3 \\ 3 \end{bmatrix}$ $(n \geqq 0)$,

$$\begin{bmatrix} 3^n,3 \\ 3,3 \\ 3 \\ 3 \end{bmatrix} (n \geqq 0), \quad [\infty], \quad [3,6], \quad [3,3,4,3], \quad \begin{bmatrix} 3,3 \\ 3,3 \\ 3,3 \end{bmatrix}, \quad \begin{bmatrix} 3,3,3 \\ 3,3,3 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 3,3,3,3,3 \\ 3,3 \\ 3 \end{bmatrix}.$$

LEMMA 9.1.[26] *If all diagonal minors of a determinant are positive, while all non-diagonal elements are negative or zero, then all algebraic first minors are positive or zero.*

Let A denote the m-rowed determinant $|\,a_j^i\,|$, A_j^i the minor derived by omitting the i^{th} row and j^{th} column, A_{jk}^{hi} that derived by omitting two rows and columns, and so on. We are given

$$A_j^j > 0, \qquad A_{jk}^{jk} > 0, \qquad A_{jk\dots}^{jk\dots} > 0, \dots, a_i^i > 0, \qquad a_j^i \leqq 0 \qquad (i \neq j),$$

and we have to prove that $(-)^{j-i} A_j^i \geqq 0$.

---

[24] For the notation, see Todd **1**, 214; Coxeter **2**, 151, 147, 162. The number of dimensions of the space in which the group operates (spherical space, in the finite case) is equal to the number of digits in the symbol, except in the case of $\boxed{3^n}$, when it is $n - 1$. For a more systematic notation, see p. 618, below.

For an independent investigation of the groups $[3^n]$, $[3^n, 4]$, $[k]$, [3, 5], [3, 4, 3], [3, 3, 5], see Motzok **1**.

[25] To avoid repetition.

[26] This lemma, being of unnecessary generality, could be replaced by Stieltjes **1**.

Since we can transpose the j^{th} and m^{th} columns (provided we also transpose the j^{th} and m^{th} rows), it will be sufficient to prove that

$$(-)^{m-i} A_m^i \geqq 0 \qquad (i < m).$$

This is true when $m = 2$, since $A_2^1 = a_1^2 \leqq 0$. Using induction, we assume the lemma as applied to A_m^m, so that

$$(-)^{j-i} A_{jm}^{im} \geqq 0.$$

We now have

$$A_m^i = -\sum_{j=1}^{m-1} (-)^{j-m} a_j^m A_{mj}^{im},$$

whence

$$(-)^{m-i} A_m^i = \sum_{j=1}^{m-1} (-a_j^m)(-)^{j-i} A_{jm}^{im} \geqq 0.$$

LEMMA 9.2. *A determinant of the kind considered in Lemma* 9.1 *is definitely diminished by simultaneously increasing* $-a_j^i$ *and* $-a_i^j$ $(i \neq j)$.

Writing $A = f(a_2^1, a_1^2)$, we have

$$f(a_2^1 - \delta, a_1^2 - \epsilon) - f(a_2^1, a_1^2) = \delta A_2^1 + \epsilon A_1^2 - \delta\epsilon A_{12}^{12}.$$

If $\delta > 0$ and $\epsilon > 0$, this expression is definitely negative, since $A_2^1 \leqq 0$, $A_1^2 \leqq 0$ and $A_{12}^{12} > 0$.

The enumeration of our irreducible groups merely involves the enumeration of simplexes whose dihedral angles are submultiples of π, for which see Coxeter **2**, 136–144 (§§15.4–15.9). The only step left obscure[27] was the "Note" in the middle of §15.2. This can be derived from Lemma 9.2 by putting $a_i^i = 1$ and $a_j^i = -\cos(i\,j)$.

THEOREM 10. *In spherical or Euclidean space, the continued product of the reflections in the bounding primes of a simplex, taken in any order, is an operation which leaves no point invariant.*

The reader will have no difficulty in proving this.

DEFINITION. By the *generators* of a g.g.r., we shall mean the reflections in the bounding primes of its fundamental region.

THEOREM 11. *Of the continued products of the generators of a finite g.g.r., arranged in various orders, any two are conjugate.*

LEMMA 11.1. *If the vertices of an m-gon are numbered from* 1 *to m in any order, we can arrange the numbers in natural order by repeated application of the following operation: whenever a side of the m-gon has inconsecutive numbers at its ends, we are allowed to transpose those numbers.*

We prove this by showing that the numbers at the ends of a side can be transposed even when they are consecutive.

---

[27] I am indebted to Prof. Pólya for pointing this out.

Suppose first that the numbers in order are

$$1, 2, p_3, p_4, \ldots, p_m,$$

so that every $p_i > 2$. Interchanging 1 with p_m, then with p_{m-1}, and so on, we obtain the new order

$$p_m, 2, 1, p_3, \ldots, p_{m-1},$$

which is the same as the old except that 1 and 2 have been transposed. We may describe this process as "making 1 run round the polygon."

Using induction, let us assume that the numbers $l-1$ and l can be transposed (when associated with consecutive vertices of the polygon). Then we can transpose l and $l+1$ by making l run round the polygon. Hence, finally, we can transpose the numbers at the ends of any side whatever, and the lemma is reduced to the familiar generation of the symmetric group by transpositions.

LEMMA 11.2. *If m vertices of an $(m+1)$-gon are numbered from 1 to m in any order, while the remaining vertex is marked l' $(1 \leqq l \leqq m)$, we can derive any rearrangement of the numbers by repeated application of the operation of Lemma 11.1, regarding l' as consecutive only to l.*

We can transpose l and l' by making l' run round the polygon; and then the presence of l' will not affect the argument used in proving Lemma 11.1.

By Theorem 9, every irreducible finite g.g.r. is of the form

$$[k_1, k_2, \ldots, k_{m-1}] \quad \text{or} \quad \begin{bmatrix} 3^n \\ 3^p \\ 3 \end{bmatrix}.$$

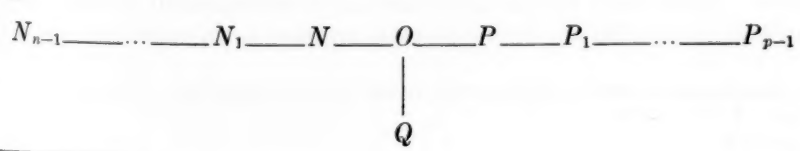
In the former case[28] the abstract definition is

$$R_i^2 = (R_i R_{i+1})^{ki} = (R_i R_j)^2 = 1 \qquad (i < j-1).$$

Thus inconsecutive R's are permutable.

Consider any ordering of the m R's. By numbering the vertices of an m-gon in accordance with the suffixes, we obtain a representation, not only of the continued product corresponding to the chosen ordering, but also of the obvious conjugates of this product, derived by cyclic permutation of the factors. The operation described in Lemma 11.1 will not affect this set of m conjugate products. Hence every such product is conjugate to $R_1 R_2 \ldots R_m$.

$\begin{bmatrix} 3^n \\ 3^p \\ 3 \end{bmatrix}$ has an abstract definition conveniently represented by the diagram



[28] Todd **1**, 224 (4).

of which the case $n = p = 2$ is explained in Coxeter **2, 164.**[29]

By associating the generators

$$N_{n-1}, \ldots, N_1, N, O, P, P_1, \ldots, P_{p-1}, Q$$

with the numbers

$$1, \ldots, l-2, l-1, l, l+1, l+2, \ldots, m, l',$$

we are able to apply Lemma 11.2, and to conclude that all continued products of these generators are conjugate.

Theorem 11 now follows, by Lemma 4.4.

These arguments can be extended to cover most of our *infinite* groups; but it is easily seen that the operations $R_1 R_2 R_3 R_4$ and $R_1 R_3 R_2 R_4$ of $\boxed{3^4}$ are *not* conjugate.

THEOREM 12. *The continued product of the generators of an infinite g.g.r., arranged in any order, is of infinite period.*

This follows easily from Lemma 4.1, Theorem 10 and Theorem 6.

DEFINITION. R (without a suffix) will denote a particular continued product of the generators of an irreducible finite g.g.r., *viz.*

$$R = R_1 R_2 \ldots R_m \quad \text{for} \quad [k_1, k_2, \ldots, k_{m-1}]$$

and

$$R = N_{n-1} \ldots N_1 N O P P_1 \ldots P_{p-1} Q \quad \text{for} \quad \begin{bmatrix} 3^n \\ 3^p \\ 3 \end{bmatrix}.$$

(We now make a brief digression into the theory of regular polytopes.)

DEFINITION. The *Petrie polygon* of a regular polygon is the regular polygon itself. The Petrie polygon of a regular polyhedron is a skew polygon of which every two consecutive sides, but no three, belong to one face.[30] The Petrie polygon of a regular polytope in m dimensions is a skew polygon of which every $m-1$ consecutive sides, but no m consecutive sides, belong to the Petrie polygon of one bounding figure. Thus, if

$$\ldots A_0 A_1 \ldots A_{m-1} A_m \ldots$$

is a Petrie polygon of the regular polytope Π_m, then

$$A_0 A_1 \ldots A_{m-1} \text{ and } A_1 \ldots A_{m-1} A_m$$

occur in Petrie polygons of two adjacent Π_{m-1}'s.

Remark. Every finite regular polytope can be orthogonally projected on to a plane in such a way that one Petrie polygon appears as an ordinary regular poly-

[29] All the generators are involutory. Pairs not directly linked are permutable. If X, Y are linked, $XYX = YXY$.

[30] Coxeter **3.**

gon whose interior is the projection of the rest of the polytope. The square aspect of the regular tetrahedron, the hexagonal aspects of the octahedron and cube, and the decagonal aspects of the icosahedron and dodecahedron, are familiar. In the projections of the four-dimensional polytopes, a second Petrie polygon (really congruent to the first, of course) appears as a star polygon,[31] *viz.*

$\{\frac{5}{2}\}$ for $\{3, 3, 3\}$, $\{\frac{8}{3}\}$ for $\{3, 3, 4\}$ and $\{4, 3, 3\}$,[32]

$\{\frac{12}{5}\}$ for $\{3, 4, 3\}$,[33] $\{\frac{30}{11}\}$ for $\{3, 3, 5\}$[34] and $\{5, 3, 3\}$.

THEOREM 13. *Of the cycles in which the operation R of $[k_1, k_2, \ldots, k_{m-1}]$ permutes the vertices of $\{k_1, k_2, \ldots, k_{m-1}\}$, one is the cycle of vertices of a Petrie polygon.*

Let X_0 be an alternative symbol for the vertex A_0, and let X_1 denote the mid-point of the edge A_0A_1, X_2 the center of the plane face $\ldots A_0A_1A_2$, X_3 the center of the solid whose Petrie polygon involves $A_0A_1A_2A_3, \ldots$ and X_m the center of the whole polytope.[35] We take R_i $(i \leqq m)$ to be the reflection in the prime determined by the m points that we obtain by omitting X_{i-1} from the set of X's.

When $m = 2$, R_1 and R_2 are the reflections in an in-radius X_1X_2 and a circum-radius A_0X_2 of the polygon $\{k\}$, X_1 being the mid-point of the side A_0A_1. Since the angle between these radii is π/k, the product $R \equiv R_1R_2$ is a rotation through $2\pi/k$, which effects a cyclic permutation of the vertices of $\{k\}$.

We thus have a basis for induction, and can assume that

$$r \equiv R_1R_2 \ldots R_{m-1}$$

effects a cyclic permutation of the vertices of the Petrie polygon

$$A_0A_1 \ldots A_{m-1}A_m' \ldots$$

of the Π_{m-1} whose center is X_{m-1}. In virtue of this inductive assumption,

$$A_{i+1} = rA_i \qquad (i < m - 1)$$

and

$$A_m' = rA_{m-1}.$$

Also, since R_m is the reflection in $A_0A_1 \ldots A_{m-2}X_m$,

$$A_i = R_mA_i \qquad (i < m - 1).$$

Hence

$$A_{i+1} = rR_mA_i$$

$$= RA_i \qquad (i < m - 1).$$

---

[31] Called $\{h'\}$ in Coxeter **3.**

[32] Van Oss **1,** Tafel I.

[33] *Ibid.*, Tafel II.

[34] *Ibid.*, Tafel VIII^a. The projections given in Tafeln VI^a, VII^a exhibit Petrie polygons of $\{5, 3, \frac{5}{2}\}$, $\{3, 5, \frac{5}{2}\}$.

[35] Coxeter **2,** 155.

Finally, since the reflection in the prime $A_1 \ldots A_{m-1}X_m$ is rR_mr^{-1},

$$\begin{aligned} A_m &= rR_mr^{-1}A_m' \\ &= rR_mA_{m-1} \\ &= RA_{m-1}. \end{aligned}$$

DEFINITION. S will denote the *central inversion, i.e.* that transformation in spherical space which replaces every point by its antipodes. S has the peculiarity of being permutable with every transformation in the spherical space.

THEOREM 14. *If the fundamental region of a finite g.g.r. has no internal symmetry, the central inversion is an operation of the group.*

Consider a particular fundamental region F. By producing its bounding primes, we see that one of the congruent regions is antipodal to F. If S' is that operation of the group which transforms F into this antipodal region, it is clear that the product $S'^{-1}S$ leaves F invariant. If F has no internal symmetry, this operation must be identity, and so $S' = S$.

COROLLARY. *If the fundamental region has internal symmetries, we may augment the g.g.r. by adjoining these symmetries. The augmented group will contain the central inversion (although the g.g.r. may or may not).*

DEFINITION. By Theorems 11 and 12, (in any discrete g.g.r.) the period of a continued product of the generators is independent of their order of arrangement. We shall call this period h. In particular, R is of period h.

THEOREM 15. *If any irreducible finite g.g.r. contains the central inversion S, then h is even and*

$$S = R^{\frac{1}{2}h}.$$

(By Theorem 11, this result would not be affected by rearranging the factors of R.)

We shall prove this theorem by examining, in each of the ten cases, the effect of the operation R on the simplest polytope whose group of symmetries includes the irreducible group under consideration. This method has the advantage that it enables us to find the actual values of h, and to see which of our groups do not contain the central inversion. (The results of this investigation are summarized in the Table on p. 618, below.)

(i) In the case of $[3^n]$, the polytope is the *regular simplex* $\{3^n\}$ or α_{n+1}. Since its vertices do not occur in opposite pairs, the group cannot contain the central inversion.

If $a_1, a_2, \ldots, a_{n+2}$ are the vertices, the generators are the transpositions

$$R_i = (a_ia_{i+1}),$$

and $R \equiv R_1R_2 \ldots R_{n+1}$ permutes the a's cyclicly.[36] Therefore

$$h = n + 2.$$

[36] Cf. Todd **1**, 229.

An $\{n+2\}$ with all pairs of vertices joined can be regarded as a projection (of the vertices and edges) of α_{n+1}. The $\{n+2\}$ itself is then the projection of a Petrie polygon.

(ii) In the case of $[3^n, 4]$, we consider the *cross polytope* $\{3^n, 4\}$ or β_{n+2}, whose vertices

$$a_1, a_2, \ldots, a_{n+2}, c_1, c_2, \ldots, c_{n+2}$$

occur in opposite pairs a_i, c_i. The generators can be taken to be reflections in the primes

$$x_1 = x_2, \quad x_2 = x_3, \ldots, \quad x_{n+1} = x_{n+2}, \quad x_{n+2} = 0,$$

viz.[37]

$$R_1 = (a_1a_2)(c_1c_2), \quad R_2 = (a_2a_3)(c_2c_3), \ldots, \quad R_{n+1} = (a_{n+1}a_{n+2})(c_{n+1}c_{n+2}),$$
$$R_{n+2} = (a_{n+2}c_{n+2}).$$

We now have

$$R = R_1R_2 \ldots R_{n+2} = (a_1a_2 \ldots a_{n+2})(c_1c_2 \ldots c_{n+2}a_{n+2})$$
$$= (a_1a_2 \ldots a_{n+2}c_1c_2 \ldots c_{n+2}).$$

Therefore $$h = 2(n+2)$$

and $$R^{\frac{1}{2}h} = (a_1c_1)(a_2c_2) \ldots (a_{n+2}c_{n+2}) = S.$$

A $\{2(n+2)\}$ with all pairs of vertices joined, except those that are diametrically opposite, can be regarded as a projection (of the vertices and edges) of β_{n+2}. The $\{2(n+2)\}$ itself is then the projection of a Petrie polygon.

(iii) In the case of $\begin{bmatrix} 3^n \\ 3 \\ 3 \end{bmatrix}$, we consider the cross polytope $\left\{3^n, \begin{matrix} 3 \\ 3 \end{matrix}\right\}$ or β_{n+3}. The generators, being reflections in the primes

$$x_1 = x_2, \quad x_2 = x_3, \ldots, \quad x_n = x_{n+1}, \quad x_{n+1} = x_{n+2}, \quad x_{n+2} = x_{n+3},$$
$$x_{n+2} + x_{n+3} = 0,$$

are

$$N_{n-1} = (a_1a_2)(c_1c_2), \quad N_{n-2} = (a_2a_3)(c_2c_3), \ldots, \quad N = (a_na_{n+1})(c_nc_{n+1}),$$
$$O = (a_{n+1}a_{n+2})(c_{n+1}c_{n+2}), \quad P = (a_{n+2}a_{n+3})(c_{n+2}c_{n+3}),$$
$$Q = (a_{n+2}c_{n+3})(c_{n+2}a_{n+3}).$$

Since $PQ = (a_{n+2}c_{n+2})(a_{n+3}c_{n+3})$, we have

$$R \equiv N_{n-1}N_{n-2} \ldots NOPQ = (a_1a_2 \ldots a_{n+2}c_1c_2 \ldots c_{n+2})(a_{n+3}c_{n+3}).$$

[37] Todd's R's are in the reverse order.

Therefore $h = 2(n + 2)$ again.

If n is even, the central inversion does not belong to the group, since the polytope $\left\{3, \begin{matrix} 3^n \\ 3 \end{matrix}\right\}$ or $h\gamma_{n+3}$ then has no opposite vertices. But if n is odd,

$$R^{n+2} = (a_1c_1)(a_2c_2) \ldots (a_{n+3}c_{n+3}) = S.$$

(iv) In the case of $[k]$, $R = R_1R_2$ and $h = k$. If k is odd, the central inversion does not belong to the group, since $\{k\}$ then has no opposite vertices. But if k is even,

$$R^{\frac{1}{2}k}$$

is the rotation through π, which is S.

(v) In the decagonal projection of the icosahedron $\{3, 5\}$, the decagon clearly represents a Petrie polygon, and opposite vertices of the Petrie polygon are opposite vertices of $\{3, 5\}$. Therefore,[38] for $[3, 5]$,

$$h = 10 \quad \text{and} \quad R^5 = S.$$

(vi) In Van Oss's dodecagonal projection of $\{3, 4, 3\}$, the peripheral $\{12\}$ represents a Petrie polygon (since four consecutive vertices belong to the Petrie polygon of a bounding octahedron), and opposite vertices of the Petrie polygon are opposite vertices of $\{3, 4, 3\}$. Therefore,[39] for $[3, 4, 3]$,

$$h = 12 \quad \text{and} \quad R^6 = S.$$

(vii) In Van Oss's triacontagonal projection of $\{3, 3, 5\}$, the peripheral $\{30\}$ represents a Petrie polygon (since four consecutive vertices belong to a bounding tetrahedron), and opposite vertices of the Petrie polygon are opposite vertices of $\{3, 3, 5\}$. Therefore,[38] for $[3, 3, 5]$,

$$h = 30 \quad \text{and} \quad R^{15} = S.$$

(viii) In the case of $\begin{bmatrix} 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$, we consider the vertices[40]

$$a_1, \cdots, a_6; \qquad b_1, \cdots, b_6; \qquad c_{12}, \cdots, c_{56}$$

of 2_{21}, and observe[41] that

$$R \equiv N_1NOPP_1Q = (654321)\,[456.123]$$

$$= (a_1a_6c_{34}c_{23}b_6b_5b_4b_3c_{16}c_{56}a_3a_2)\ (a_5c_{35}c_{24}c_{13}b_1c_{12}b_2c_{26}c_{15}c_{46}a_4c_{45})\ (c_{36}c_{25}c_{14}),$$

whence $$h = 12.$$

---

[38] Todd **1**, 230.

[39] *Ibid.*, 227.

[40] Coxeter **1**, 388 (§9.4). Todd **2** provided the most satisfactory proof of the correspondence between these vertices and the lines on the general cubic surface (Schläfli **1**, Dickson **1**).

[41] Coxeter **2**, 165 (18.31).

Since 2_{21} has no opposite vertices, the central inversion is absent.

(ix) In the case of $\begin{bmatrix} 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$, we consider the vertices[42]

$$c_{12}, \cdots, c_{78}; \qquad C_{12}, \cdots, C_{78}$$

of 3_{21}, and observe[43] that

$$R \equiv N_2N_1NOPP_1Q = (7654321)\,[4567.1238]$$

$$= (C_{17}C_{67}c_{34}c_{23}C_{78}C_{68}C_{58}C_{48}C_{38}c_{17}c_{67}C_{34}C_{23}c_{78}c_{68}c_{58}c_{48}c_{38})$$
$$(C_{57}c_{35}c_{24}c_{13}C_{18}c_{12}C_{28}c_{27}c_{16}c_{57}C_{35}C_{24}C_{13}c_{18}C_{12}c_{28}C_{27}C_{16})$$
$$(c_{36}c_{25}c_{14}c_{37}c_{26}c_{15}c_{47}C_{45}c_{56}C_{36}C_{25}C_{14}C_{37}C_{26}C_{15}C_{47}c_{45}C_{56})$$
$$(c_{46}C_{46}),$$

whence $$h = 18 \quad \text{and} \quad R^9 = (cC) = S.$$

(x) In the case of $\begin{bmatrix} 3, 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$, we consider the vertices[44]

$$a_{123}, \cdots, a_{789}; \qquad b_{123}, \cdots, b_{789}; \qquad c_{12}, \cdots, c_{98}$$

of 4_{21}, and observe[45] that

$$R \equiv N_3N_2N_1NOPP_1Q = (87654321)\,V_{123}$$

$$= (a_{178}a_{678}b_{349}b_{239}c_{98}c_{97}c_{96}c_{95}c_{94}c_{93}b_{189}b_{789}a_{345}a_{234}c_{83}$$
$$b_{178}b_{678}a_{349}a_{239}c_{89}c_{79}c_{69}c_{59}c_{49}c_{39}a_{189}a_{789}b_{345}b_{234}c_{38})$$
$$(a_{578}b_{359}b_{249}b_{139}c_{91}b_{129}c_{92}b_{289}b_{179}b_{689}a_{346}a_{235}c_{84}c_{73}b_{168}$$
$$b_{578}a_{359}a_{249}a_{139}c_{19}a_{129}c_{29}a_{289}a_{179}a_{689}b_{346}b_{235}c_{48}c_{37}a_{168})$$
$$(a_{369}a_{259}a_{149}a_{389}a_{279}a_{169}a_{589}b_{356}b_{245}b_{134}c_{31}c_{28}a_{278}a_{167}a_{568}$$
$$b_{369}b_{259}b_{149}b_{389}b_{279}b_{169}b_{589}a_{356}a_{245}a_{134}c_{13}c_{82}b_{278}b_{167}b_{568})$$
$$(a_{379}a_{269}a_{159}a_{489}b_{456}a_{679}b_{347}b_{236}c_{58}c_{47}c_{36}a_{158}a_{478}b_{459}a_{567}$$
$$b_{379}b_{269}b_{159}b_{489}a_{456}b_{679}a_{347}a_{236}c_{85}c_{74}c_{63}b_{158}b_{478}a_{459}b_{567})$$
$$(a_{148}a_{378}a_{267}a_{156}a_{458}b_{569}a_{367}a_{256}a_{145}a_{348}a_{237}c_{86}c_{75}c_{64}c_{53}$$
$$b_{148}b_{378}b_{267}b_{156}b_{458}a_{569}b_{367}b_{256}b_{145}b_{348}b_{237}c_{68}c_{57}c_{46}c_{35})$$
$$(a_{468}b_{469}a_{467}b_{479}a_{457}b_{579}a_{357}a_{246}a_{135}c_{14}a_{123}b_{128}c_{72}b_{268}b_{157}$$
$$b_{468}a_{469}b_{467}a_{479}b_{457}a_{579}b_{357}b_{246}b_{135}c_{41}b_{123}a_{128}c_{27}a_{268}a_{157})$$
$$(a_{368}a_{257}a_{146}a_{358}a_{247}a_{136}c_{15}a_{124}c_{23}c_{12}c_{81}b_{127}c_{62}b_{258}b_{147}$$
$$b_{368}b_{257}b_{146}b_{358}b_{247}b_{136}c_{51}b_{124}c_{32}c_{21}c_{18}a_{127}c_{26}a_{258}a_{147})$$
$$(c_{17}a_{126}c_{25}a_{248}a_{137}c_{16}a_{125}c_{24}a_{238}c_{87}c_{76}c_{65}c_{54}c_{43}b_{138}$$
$$c_{71}b_{126}c_{52}b_{248}b_{137}c_{61}b_{125}c_{42}b_{238}c_{78}c_{67}c_{56}c_{45}c_{34}a_{138}),$$

whence $$h = 30 \quad \text{and} \quad R^{15} = (ab)(c_{ij}c_{ji}) = S.$$

---

[42] Coxeter **1**, 387.
[43] Coxeter **2**, 168.
[44] *Ibid.*, 170.
[45] *Ibid.*, 172.

THEOREM 16. *If an irreducible discrete g.g.r. operates in Euclidean space of m dimensions (or spherical space of m − 1 dimensions), the total number of reflections is* $\frac{1}{2}mh$.

LEMMA 16.1. *In an irreducible discrete g.g.r. whose symbol contains no even digits, all reflections are conjugate.*

Since every reflection in the group is conjugate to some generator, it will be sufficient to prove that the generators are conjugate. If U, Z are any two generators, our assumption implies that we can find a sequence of generators $U, V, W, \cdots, Y, Z$, such that the products $UV, VW, \cdots, YZ$ are all of odd period. Suppose

$$(UV)^{2i+1} = 1.$$

Then

$$(UV)^i U (VU)^i = V,$$

whence U and V are conjugate. Proceeding similarly throughout the sequence, we conclude that U and Z are conjugate.

For infinite groups, Theorem 16 is obvious, since h is infinite. For finite groups, we consider the actual cases (as in Theorem 15).

(i)[46] $[3^n]$. $m = n + 1, \quad h = m + 1.$

The transposition $(a_i a_j)$ is the reflection that reverses the edge $a_i a_j$ of α_m, leaving the opposite α_{m-2} invariant. By Lemma 16.1, all the reflections are of this kind. Hence their number is $\frac{1}{2}m(m+1)$.

(ii)[47] $[3^n, 4]$. $m = n + 2, \quad h = 2m.$

Since

$$(R_1R_2)^3 = (R_2R_3)^3 = \cdots = (R_{m-2}R_{m-1})^3 = 1,$$

every reflection is conjugate either to R_1 or to R_m. R_1 reverses each of the opposite edges a_1a_2, c_1c_2 of β_m; R_m transposes the opposite vertices a_m, c_m. Since β_m has $2m(m-1)$ edges and $2m$ vertices, the total number of reflections is

$$m(m-1) + m = m^2.$$

(iii)[48] $\begin{bmatrix} 3^n \\ 3 \\ 3 \end{bmatrix}$. $m = n + 3, \quad h = 2(m-1).$

By Lemma 16.1, every reflection is conjugate to that which reverses the opposite edges a_1a_2, c_1c_2 of β_m. Hence the number of reflections is $m(m-1)$.

---

[46] For the case when $n = 3$, cf. Goursat **1**, 90.
[47] For the case when $n = 2$, *ibid.*, 85.
[48] For the case when $n = 1$, *ibid.*, 82.

(These same reflections occur in $[3^{m-2}, 4]$, whose remaining m reflections generate $[\,]^m$.)[49]

(iv) $[k]. \quad m = 2, \quad h = k.$

If k is odd, every reflection reverses one side of $\{k\}$ and leaves the opposite vertex invariant; so the number is k.

If k is even, R_1 reverses each of two opposite sides of $\{k\}$, while R_2 leaves two opposite vertices invariant; so the number is $\frac{1}{2}k + \frac{1}{2}k$.[50] (Although R_1 and R_2 are not conjugate in $[k]$, these same operations are conjugate in the larger group $[2k]$.)

(v) $[3, 5]. \quad m = 3, \quad h = 10.$

Every plane of symmetry of $\{3, 5\}$ passes through one pair of opposite edges, and bisects another pair; so the number of planes is 15.

(vi)[51] $[3, 4, 3]. \quad m = 4, \quad h = 12.$

Since

$$(R_1R_2)^3 = (R_3R_4)^3 = 1,$$

there are only two sets of mutually conjugate reflections. But there is a reflection that transposes two opposite vertices of $\{3, 4, 3\}$, and another that transposes two opposite octahedra. Hence the total number is $12 + 12$.[52] (*All* these reflections are conjugate in the group of symmetries of $t_{1,2}\{3, 4, 3\}$.[53])

(vii)[54] $[3, 3, 5]. \quad m = 4, \quad h = 30.$

There is a reflection that transposes two opposite vertices of $\{3, 3, 5\}$. By Lemma 16.1, every reflection is of this kind. Hence the number is 60.

(viii) $\begin{bmatrix} 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}. \quad m = 6, \quad h = 12.$

---

[49] $[]^m$ means $[] \times [] \times \cdots$ (to m terms), which might also be written $[2^{m-1}]$. (See *Preface* for the cases $m = 2, 3$.)

The relation between these groups is epitomized in the symbolic equation

$$[3^{m-2}, 4] = \begin{bmatrix} 3^{m-3} \\ 3 \\ 3 \end{bmatrix} + []^m \ (m \geqq 3).$$

[50] $[k] = [\frac{1}{2}k] + [\frac{1}{2}k]$ (k even).

[51] Goursat **1**, 86.

[52] $[3, 4, 3] = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$

[53] This is an example of the "augmented group" defined in the corollary to Theorem 14. It is derived from [3, 4, 3] by adjoining the operation considered in Robinson **1**.

[54] Goursat **1**, 88.

P_1 transposes the two opposite α_5's[55]

$$a_1b_1c_{23}c_{24}c_{25}c_{26}, \qquad a_2b_2c_{13}c_{14}c_{15}c_{16}$$

of 2_{21}. Since there are 72 α_5's, the number of reflections is 36. (These correspond, of course, to the double-sixes on the cubic surface.)

(ix) $\begin{bmatrix} 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$. $\quad m = 7, \quad h = 18.$

P_1 transposes the two opposite β_6's

$$c_{13} \cdots c_{18}C_{23} \cdots C_{28}, \qquad C_{13} \cdots C_{18}c_{23} \cdots c_{28}$$

of 3_{21}. Since there are 126 β_6's, the number of reflections is 63. (These correspond[56] to the 63 *period characteristics* of genus 3.)

(x) $\begin{bmatrix} 3, 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$. $\quad m = 8, \quad h = 30.$

P_1 transposes the two opposite vertices c_{12}, c_{21} of 4_{21}. Since there are 240 vertices, the number of reflections is 120.[57]

THEOREM 17. *Every irreducible finite g.g.r. can be generated by two operations.*

For all finite groups of the form $[k_1, k_2, \cdots, k_{m-1}]$, this was proved by Todd,[58] the two generating operations being, except in one case, R_1 and R.

In the case of $[3^n]$, putting $R = R_1R_2 \cdots R_{n+1}$, we have

$$R_{i+1} = R^iR_1R^{-i} \qquad (i \leqq n).$$

In the case of $[3^n, k]$, putting $R = R_1R_2 \cdots R_{n+2}$, we have

$$R_{i+1} = R^iR_1R^{-i} \qquad (i \leqq n),$$

$$R_{n+2} = R^nR_1(R^{-1}R_1)^nR.$$

This result includes $[3^n, 4]$, $[k]$, [3, 5], [3, 3, 5], and incidentally also the infinite groups $[\infty]$, [3, 6]. Todd[59] generates the exceptional group [3, 4, 3], by means of $R_1R_2R_4$ and $R_1R_3R_4$.

In the case of $\begin{bmatrix} 3^n \\ 3 \\ 3 \end{bmatrix}$, we define

$$U \equiv N_{n-1}N_{n-2} \cdots NO, \qquad V \equiv POQ,$$

---

[55] Coxeter **2**, 165.
[56] Du Val **1**, 58.
[57] *Ibid.*, 47, 49.
[58] Todd **1**, 226.
[59] *Ibid.*, 228.

and observe that

$$\begin{aligned} U^{-1}OU &= ONONO \qquad \text{(since } O \text{ is permutable with } N_1, \cdots, N_{n-1}\text{)} \\ &= N, \\ U^{-1}N_iU &= ON \cdots N_{i-1}N_iN_{i+1}N_iN_{i+1}N_iN_{i-1} \cdots NO \\ &= ON \cdots N_{i-1}N_{i+1}N_{i-1} \cdots NO \\ &= N_{i+1}, \end{aligned}$$

$$V^{-1}OV = QPQ = P, \qquad VOV^{-1} = PQP = Q,$$

and

$$U^{-1}VU = ONPOQNO = OPNONQO = OPONOQO = POPNQOQ$$
$$= POQNPOQ = VNV.$$

Since $N = N^{-1}$, we now have

$$\begin{aligned} N &= VU^{-1}V^{-1}UV, \\ N_i &= U^{-i}NU^i, \\ O &= UNU^{-1}, \\ P &= V^{-1}UNU^{-1}V, \\ Q &= VUNU^{-1}V^{-1}. \end{aligned}$$

When n is even, we can alternatively generate the group by means of O and

$$R' \equiv N_{n-1}N_{n-2} \cdots NPOQ.^{60}$$

In fact,

$$\begin{aligned} N &= OR'^{n+1}OR'^{-n+1}O, \\ N_i &= R'^{-i}NR'^i, \\ P &= R'^{-(n+1)}OR'^{n+1}, \\ Q &= R'OR'^{-1} \qquad (n \text{ even}^{61}). \end{aligned}$$

The remaining groups are the symmetry groups of the polytopes 2_{21}, 3_{21}, 4_{21}, which were considered in Coxeter **1**.[62] However, we shall reconsider them here, in order to show that they can all be generated by the special operations O, R.

---

[60] By Theorem 11, R' is conjugate to R (which has OP in place of PO). But the permutable operations P, Q could never be distinguished in terms of O and R.

[61] When n is odd, $R'^{n+2} = S$, so that

$$R'^{n+1}OR'^{-(n+1)} = R'^{-1}OR' = QPNONPQ$$

and

$$R'^{-(n+1)}OR'^{n+1} = R'OR'^{-1} = Q.$$

[62] 408 (11.63), 406 (11.52), 409 (11.71).

We can express the generators of $\begin{bmatrix} 3^n \\ 3^2 \\ 3 \end{bmatrix}$ in terms of O and

$$R \equiv N_{n-1}N_{n-2}\cdots NOPP_1Q,$$

by observing that

$$\begin{aligned}
R^{-1}OR &= QP_1PONONOPP_1Q = QP_1PNPP_1Q \\
&= N, \\
R^{-1}N_iR &= QP_1PONN_1\cdots N_{i-1}N_iN_{i+1}N_iN_{i+1}N_iN_{i-1}\cdots N_1NOPP_1Q \\
&= QP_1PONN_1\cdots N_{i-1}N_{i+1}N_{i-1}\cdots N_1NOPP_1Q \\
&= N_{i+1}, \\
R^{-2}N_{n-1}R^2 &= QP_1PONN_1\cdots N_{n-2}N_{n-1}QP_1PONN_1\cdots N_{n-2} \\
&\qquad N_{n-1}N_{n-2}\cdots N_1NOPP_1QN_{n-1}N_{n-2}\cdots N_1NOPP_1Q \\
&= QP_1PONN_1\cdots N_{n-2}QP_1PONN_1\cdots N_{n-3}N_{n-1}N_{n-2} \\
&\qquad N_{n-1}N_{n-2}N_{n-1}N_{n-3}\cdots N_1NOPP_1QN_{n-2}\cdots N_1NOPP_1Q \\
&= QP_1PONN_1\cdots N_{n-2}QP_1PONN_1\cdots N_{n-3} \\
&\qquad N_{n-2}N_{n-3}\cdots N_1NOPP_1QN_{n-2}\cdots N_1NOPP_1Q \\
&= \cdots \\
&= QP_1PONQP_1PONOPP_1QNOPP_1Q \\
&= P_1PP_1QOQPNONONPQOQP_1PP_1 \\
&= PP_1POQOPOPOQOPP_1P = PP_1POPOPP_1P = POP \\
&= OPO, \\
RPR^{-1} &= N_{n-1}N_{n-2}\cdots NOPP_1PP_1PON\cdots N_{n-2}N_{n-1} \\
&= P_1, \\
R^{-1}PR &= QP_1POPOPP_1Q = QOQ \\
&= OQO,
\end{aligned}$$

whence

$$(17.1)\qquad \begin{cases} N_{i-1} = R^{-i}OR^i & (N = N_0), \\ P = OR^{-(n+2)}OR^{n+2}O, \\ P_1 = ROR^{-(n+2)}OR^{n+2}OR^{-1}, \\ Q = OR^{-1}OR^{-(n+2)}OR^{n+2}ORO. \end{cases}$$

These relations can be simplified considerably when $n = 2$, and slightly when $n = 3$.

For $\begin{bmatrix} 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$ $(n = 2, R = N_1 NOPP_1 Q)$, we find

$$R^3 OR^{-3} = OQO,$$

so that

$$N_1 = R^{-2} OR^2,$$
$$N = R^{-1} OR,$$
$$P = R^4 OR^{-4},$$
$$P_1 = R^5 OR^{-5},$$
$$Q = OR^3 OR^{-3} O.$$

For $\begin{bmatrix} 3, 3, 3 \\ 3, 3, \\ 3, \end{bmatrix}$ $(n = 3, R = N_2 N_1 NOPP_1 Q)$,

$$R^9 = S,$$

so that

$$R^{-5} OR^5 = R^4 OR^{-4}$$

and

$$N_2 = R^{-3} OR^3,$$
$$N_1 = R^{-2} OR^2,$$
$$N = R^{-1} OR,$$
$$P = OR^4 OR^{-4} O,$$
$$P_1 = ROR^4 OR^{-4} OR^{-1},$$
$$Q = OR^{-1} OR^4 OR^{-4} ORO.$$

For $\begin{bmatrix} 3, 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$, we cannot do better than put $n = 4$ in (17.1). Incidentally, by putting $n = 5$ we obtain a generation of the *infinite* group $\begin{bmatrix} 3, 3, 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$ by two[63] operations. For $n > 5$ we have the (infinite) symmetry groups of Du Val's Minkowskian polytopes[64] n_{21}.

[63] Three were used in Coxeter **1**, 409.
[64] Du Val **1**, 71.

DEFINITION. If Γ is a group which contains the central inversion S,

$$\tfrac{1}{2}\Gamma$$

will denote the factor group Γ/Σ, where Σ is the subgroup of order 2 generated by S. From an abstract definition for Γ, we derive an abstract definition for $\frac{1}{2}\Gamma$ by inserting the new relation $S = 1$.

Examples.

(i) $\frac{1}{2}[2k] \sim [k]$, both having the abstract definition

$$R_1^2 = R_2^2 = (R_1R_2)^k = 1.$$

(ii) $\frac{1}{2}[3, 4] \sim [3, 3]$.

$\frac{1}{2}[3, 4]$[65] has the abstract definition

$$R_1^2 = R_2^2 = R_3^2 = (R_1R_2)^3 = (R_2R_3)^4 = (R_3R_1)^2 = (R_2R_3R_1)^3 = 1.$$

On putting $R_3 = R_3'R_1$, this becomes

$$R_1^2 = R_2^2 = R_3'^2 = (R_1R_2)^3 = (R_2R_3')^3 = (R_3'R_1)^2 = (R_2R_3'R_1)^4 = 1,$$

which is the abstract definition for [3, 3] (with the superfluous relation $R^4 = 1$).

(iii) In the notation of Coxeter **2**, 145–151, (since the central inversion is a negative operation when the number of Euclidean dimensions is odd)

$$\tfrac{1}{2}[3^n, 4] \sim [3^n, 4]' \qquad (n \text{ odd}),$$

$$\tfrac{1}{2}[3, 5] \sim [3, 5]' \qquad (\text{the icosahedral group}^{66}),$$

$$\tfrac{1}{2}\begin{bmatrix} 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix} \sim \begin{bmatrix} 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}'.$$

Du Val **1**, 58 shows that the last of these groups is the θ-characteristic group of genus 3. We now see that an abstract definition for it can be obtained by adjoining the relation

$$R^9 = 1$$

to the abstract definition[67] for $\begin{bmatrix} 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$. This new definition is more elegant than (18.51).[68] In terms of the digits $1, 2, \cdots, 8$, whose pairs denote the twenty-eight odd θ-characteristics (or the twenty-eight bitangents of the plane quartic curve), Q (or Q_0) is simply a bifid substitution, while the other six generators are transpositions.[69] By 15 (ix) above, R permutes twenty-seven of the

[65] Fricke-Klein **1**, 71.
[66] *Ibid.*, 72.
[67] Coxeter **2**, 149 (16.74).
[68] *Ibid.*, 166.
[69] *Ibid.*, 168.

odd characteristics (or bitangents) in three cycles, namely

(17 67 34 23 78 68 58 48 38)

(57 35 24 13 18 12 28 27 16)

(36 25 14 37 26 15 47 45 56).

(iv) $\frac{1}{2}\begin{bmatrix} 3,3,3,3 \\ 3,3 \\ 3 \end{bmatrix}$ is the group considered in Du Val **1,** 49. The "one extra relation" that was required in Coxeter **2,** 174 (at the bottom of the page) is now seen to be

$$R^{15} = 1,$$

where

$$R = N_3 N_2 N_1 NOPP_1 Q.$$

(v) Of Goursat's groups,[70]

XLII is $\frac{1}{2}\begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$,

XLIV is $\frac{1}{2}([\,] \times [3, 4])$,

XLV is $\frac{1}{2}[3, 4, 3]$,

XLVII is $\frac{1}{2}[3, 3, 4]$,

XLIX is $\frac{1}{2}([\,] \times [3, 5])$,

L is $\frac{1}{2}[3, 3, 5]$;

But LI $\sim [3, 3, 3]$ itself, the central inversion being absent.

Remark. If Γ (containing S) is a g.g.r. in spherical space, then $\frac{1}{2}\Gamma$ is a g.g.r. in *elliptic* space (of the same number of dimensions).

THEOREM 18. *For every finite g.g.r. in Euclidean (or spherical) space, there is a simply isomorphic group in elliptic space. Every other discrete g.g.r. in elliptic space is of the form*

$$\tfrac{1}{2}(\Gamma^{(1)} \times \Gamma^{(2)} \times \cdots),$$

where all the Γ*'s*[71] *occur among*

$[3^n, 4]$, $\begin{bmatrix} 3^n \\ 3 \\ 3 \end{bmatrix}$ (n odd), $[k]$ (k *even*), $[3, 5]$, $[3, 4, 3]$, $[3, 3, 5]$,

$\begin{bmatrix} 3,3,3 \\ 3,3 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 3,3,3,3 \\ 3,3 \\ 3 \end{bmatrix}$.

---

[70] Goursat **1,** 78, 79.

[71] There may, of course, be only one.

LEMMA 18.1. *If* Γ, *a finite g.g.r., contains* S, *then*

$$\tfrac{1}{2}([\,] \times \Gamma) \sim \Gamma.$$

LEMMA 18.2. *If a reducible group contains the central inversion, then each component must contain its own central inversion.*

(The proofs of these lemmas are omitted, as they involve no ingenuity.)

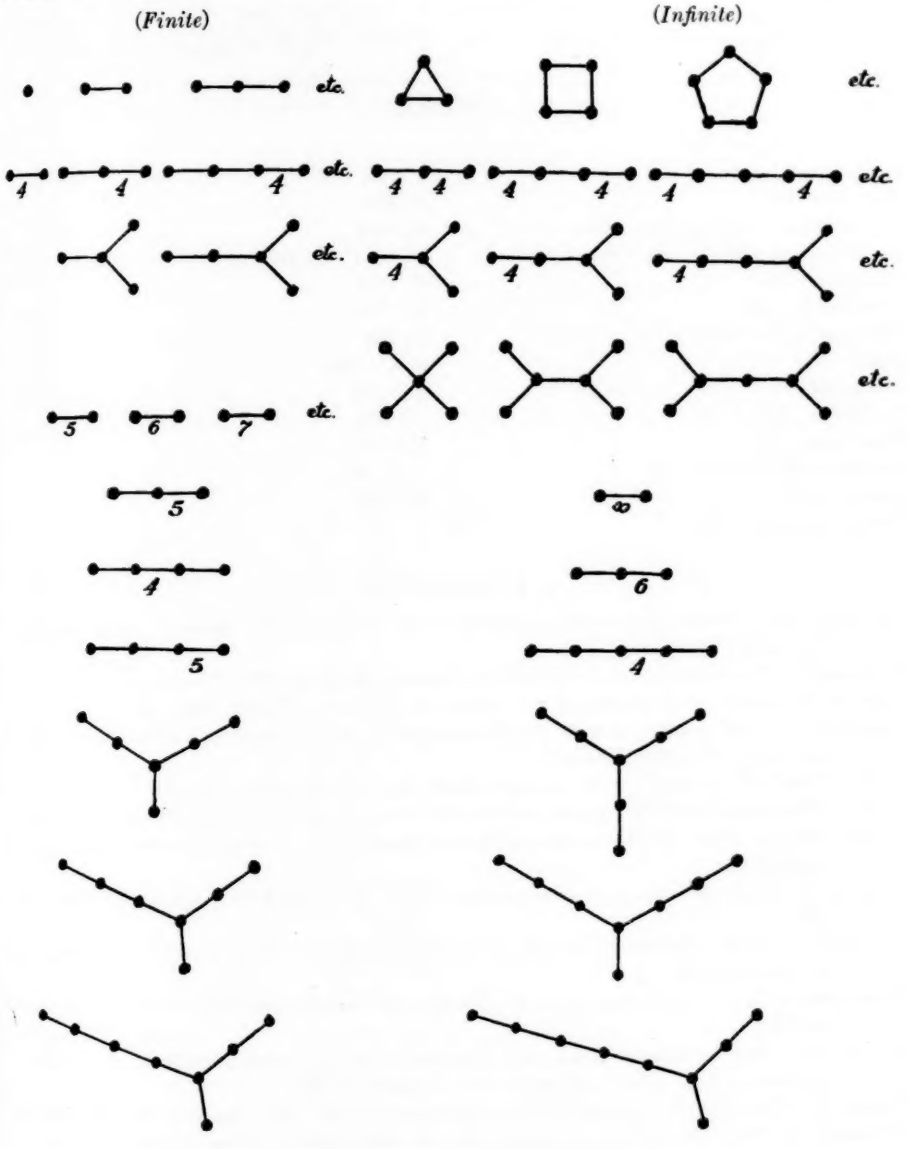
It is obvious that any group in elliptic space can be derived from a group in spherical space by identifying antipodal points. If the spherical group does not contain the central inversion, the elliptic group is simply isomorphic (but the fundamental region of the latter is one half of that of the former). If the spherical group is of the form $[\,] \times \Gamma$ and contains the central inversion, the derived elliptic group is simply isomorphic with Γ, by Lemma 18.1. The remaining case is when the spherical group contains the central inversion but no component $[\,]$ (*i.e.*, no single reflection which is permutable with all the others). By Lemma 18.2, every component of the spherical group must then contain its own central inversion.

TABLE OF IRREDUCIBLE FINITE GROUPS GENERATED BY REFLECTIONS

Group	Order[72]	m	h	Central inversion?
$[3^n]$	$(n+2)!$	$n+1$	$n+2$	Only when $n = 0$
$[3^n, 4]$	$2^{n+2}(n+2)!$	$n+2$	$2(n+2)$	Yes
$\begin{bmatrix} 3^n \\ 3 \\ 3 \end{bmatrix}$	$2^{n+2}(n+3)!$	$n+3$	$2(n+2)$	Only when n is odd
$[k]$	$2k$	2	k	Only when k is even
[3, 5]	120	3	10	Yes
[3, 4, 3]	1152	4	12	Yes
[3, 3, 5]	14400	4	30	Yes
$\begin{bmatrix} 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$	51840	6	12	No
$\begin{bmatrix} 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$	2903040	7	18	Yes
$\begin{bmatrix} 3, 3, 3, 3 \\ 3, 3 \\ 3 \end{bmatrix}$	696729600	8	30	Yes

[72] Coxeter **2**, 159.

GRAPHICAL REPRESENTATION OF THE IRREDUCIBLE GROUPS GENERATED BY REFLECTIONS



Here each dot represents a *mirror*, i.e. the prime of a generating reflection. The links indicate the relative positions of the various mirrors. When two dots are joined by a link marked k, the corresponding mirrors are inclined at angle π/k. When a link is unmarked, we understand $k = 3$. When two dots are not (directly) linked, the mirrors are perpendicular.[73]

[73] *Cf.* Coxeter 4.

By Theorem 9, the above graphs represent *all* irreducible discrete groups generated by reflections. By Theorem 6, the graph for any *reducible* discrete g.g.r. is obtained by juxtaposing several of these, repetitions being allowed.

INDEX

Central inversion: 606
Components: 595
Discrete: 594
Fricke-Klein construction: 595
Fundamental region: 594, 595
Generator: 602
Irreducible: 595
Petrie polygon: 604
Prime: 595
Prism: 591
Reducible: 595
Secundum: 595
Simplex-polycorypha: 590
Sphere: 591
Vertex diagram: 592
g.g.r.: 596
h: 606
m: 610
N, O, P, Q: 603
R: 604
R_i: 605
S: 606
$\bar{\omega}$, $\bar{\omega}'$: 595
Σ_m: 591
[ ]: 588
$[k]$: 588
$[\infty]$: 590
$\frac{1}{2}\Gamma$: 616

REFERENCES

BIEBERBACH 1: "Ueber die Bewegungsgruppen der Euklidischen Räume," *Math. Annalen*, 70 (1911), 297–336.

BURNSIDE 1: *Theory of groups of finite order* (Cambridge, 1911), 263 (*Theorem* VI).

CARTAN 1: *Leçons sur la géométrie des espaces de Riemann* (Paris, 1928), 75.

COXETER 1: "The polytopes with regular-prismatic vertex figures," *Phil. Trans. Royal Soc.* (A), 229 (1930), 329–425.

2: (Part 2 of the same), *Proc. London Math. Soc.* (2), 34 (1932), 126–189.

3: "The densities of the regular polytopes," *Proc. Camb. Phil. Soc.*, 27 (1931), 202–204.

4: "Groups whose fundamental regions are simplexes," *Journal London Math. Soc.*, 6 (1931), 133.

DICKSON 1: *Linear groups, with an exposition of the Galois field theory* (Leipzig, 1901), 303–306.

DU VAL 1: "On the directrices of a set of points in a plane," *Proc. London Math. Soc.* (2), 35 (1933), 23–74.

FRICKE-KLEIN 1: *Vorlesungen über die Theorie der automorphen Functionen*, 1 (Leipzig, 1897).

GOURSAT 1: "Sur les substitutions orthogonales et les divisions régulières de l'espace," *Annales Sci. de l'École Normale* (3), 6 (1889), 9–102.

MOTZOK 1: "The complete groups of the regular polytopes," *Mat. Sbornik*, 40 (1933), 86–114.

ROBINSON 1: "On the orthogonal groups in four dimensions," *Proc. Camb. Phil. Soc.*, 27 (1931), 43 (13).

SCHLÄFLI 1: "An attempt to determine the twenty-seven lines upon a surface of the third order ... ," *Quart. J. of Math.*, 2 (1858), 116.

SCHOENFLIES 1: *Krystallsysteme und Krystallstructur* (Leipzig, 1891), 146.

SCHOUTE 1: "The characteristic numbers of the prismotope," *Proc. Royal Acad. of Sci., Amsterdam*, 14 (1911), 424–428.

SOMMERVILLE 1: *An introduction to the geometry of n dimensions* (London, 1929), 102, 103.

STIELTJES 1: "Sur les racines de l'équation $X_n = 0$," *Oeuvres complètes*, 2 (Groningen, 1918), 75 (*Corollaire*).

TODD 1: "The groups of symmetries of the regular polytopes," *Proc. Camb. Phil. Soc.*, 27 (1931), 212–231.

2: "Polytopes associated with the general cubic surface," *Journal London Math. Soc.*, 7 (1932), 200–205.

VAN OSS 1: "Das regelmässige Sechshundertzell und seine selbstdeckenden Bewegungen," *Ver. der K. Akad. van Wet. te Amsterdam* (1), 7.1 (1899).

PRINCETON UNIVERSITY, N. J.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

COMPOSITION AND RANK OF n-WAY MATRICES AND MULTILINEAR FORMS[1]

BY RUFUS OLDENBURGER

(Received November 16, 1932, revised June 20, 1933)

1. **Introduction.** Zehfuss[2] proved that a full sign determinant of even class of a system of forms is a relative invariant of these forms, and showed that the full sign determinant of an n-adic form F of even degree is an invariant of index equal to the degree n of F. Hitchcock[3] proved that the 2-way ranks of polyadics are invariant under non-singular linear transformations. The matrices associated with polyadics possess certain factorization properties;[4] i.e., if $A = (a_{i_1 \dots i_p})$ is the matrix of a polyadic, then

$$A = \left(\sum_{j=1}^{h} a_{1ji_1} \dots a_{pji_p} \right),$$

where $(a_{1ji_1}), \cdots, (a_{pji_p})$ are 2-way matrices. Rice[5] proved certain bound properties of 2-way ranks of general n-way matrices. Hitchcock[6] further proved by polyadic methods the invariance of certain space ranks of the matrix A above under non-singular linear transformations.

The present paper begins with the study of certain operations on n-way matrices which the author has designated *compositions*. These operations include sets of linear transformations on one or more indices, Cremona transformations, the transformations of the characteristic matrices of a linear algebra, the transformations of tensors and polyadics, etc. Bound properties of 2-way ranks under compositions, relations between the ranks of a given matrix, and the canonical forms which depend on these ranks are established. Space ranks of a general signancy are introduced and several properties of these ranks are obtained. Finally, the invariance of ranks of compound matrices under non-singular compositions is proven by the use of multiple concomitants.

By means of the equivalence of a set of transformations to a single composition,

---

[1] Presented to the American Mathematical Society, December, 1930.

[2] Zehfuss, Ueber die Erweiterung des Begriffes der Determinanten, Programme Gerwerbschule Frankfort a/M. (1868), p. 6.

[3] F. L. Hitchcock, The expression of a tensor or a polyadic as a sum of products, Journal of Mathematics and Physics, vol. 6 (1927), pp. 164–189. We shall call this paper "Hitchcock, Paper 1."

[4] F. L. Hitchcock, Ordered determinants with application to polyadics, Journal of Mathematics and Physics, vol. 4 (1925), pp. 222, 223.

F. L. Hitchcock, Paper 1, p. 164.

[5] L. H. Rice, Couche ranks in the general matrix, Journal of Mathematics and Physics, vol. 7 (1928), pp. 94–96.

[6] F. L. Hitchcock, Multiple invariants and generalized rank of a p-way matrix or tensor, Journal of Mathematics and Physics, vol. 7 (1927), pp. 40–79.

the author is able to show how the complete invariant theory of 2-way matrices can be extended to n-way matrices.

Throughout the paper repetition of subscripts will denote summation.

Unless otherwise stated the proofs hold for an arbitrary field of numbers commutative under multiplication.

2. **Definitions.** It is assumed that the reader is familiar with the terminology used in standard works on ordinary (2-way) matrices. We shall also need the following terms.

A *display* $(a_{i_1 i_2 \cdots i_p})$ of a p-way matrix A is defined to be the array of numbers obtained by placing $a_{i_1 i_2 \cdots i_p}$ at the point $(i_1, i_2, \cdots i_p)$ in a linear p-space where $i_j = 1, 2, \cdots, n_j$ for $j = 1, 2, \cdots, p$. The j-th direction in this space is the direction associated with the index i_j. Other p-way displays of A are obtained by transpositions of the indices in $(a_{i_1 i_2 \cdots i_p})$.

A set of indices of A determines an *aspect*[7] or orientation in the space of A, a single index determining a direction in this space. An ordered set of indices $T = i_r i_s \cdots i_u i_w$ of the indices of A is called the *partition* T. A partition is of *class* n if it contains n indices. The class of the partition containing all of the indices of A is called the class of A. The *order* of a partition T is the product of the orders of the ranges of the indices contained in T. A partition σ will be called a *subset* of T if the indices of σ are contained in T. Two partitions T_1, T_2 will be said to be equal if they have the same ranges on corresponding indices (1st, 2nd, $\cdots$). If ρ is the partition $i_r i_s \cdots$, σ is the partition $i_q i_t \cdots$, $\cdots$, γ is the partition $i_u i_w \cdots$, where $T = i_r i_s \cdots i_q i_t \cdots i_u i_w \cdots$, then $\rho, \sigma, \cdots, \gamma$ are said to simply cover the partition T. If $\rho, \sigma, \cdots, \gamma$ simply cover T, we write $T = \rho\sigma \cdots \gamma$.

Let $T_1, T_2, \cdots, T_r$ be mutually exclusive, exhaustive partitions of the indices of A. We define the display $(a_{T_1 T_2 \cdots T_r})$ of A to be the matrix obtained by letting $T_1, \cdots, T_r$ denote the indices of an r-way matrix, indices in each partition varying by convention from right to left. For example, if $A = (a_{i_1 i_2 i_3 i_4})$, $i_1, i_2, i_3, i_4 = 1, 2$, and $T_1 = i_1 i_2$, $T_2 = i_3 i_4$, then

$$A = (a_{T_1 T_2}) = \begin{bmatrix} a_{1111} & a_{1112} & a_{1121} & a_{1122} \\ a_{1211} & a_{1212} & a_{1221} & a_{1222} \\ a_{2111} & a_{2112} & a_{2121} & a_{2122} \\ a_{2211} & a_{2212} & a_{2221} & a_{2222} \end{bmatrix}.$$

In the general case where $r = 2$, we obtain the display $(a_{T_1 T_2})$ which possesses a 2-way *determinant* $| a_{T_1 T_2} |$ on T_1, T_2 provided that the orders of T_1 and T_2 are equal. A matrix is non-singular on T_1, T_2 if $| a_{T_1 T_2} | \neq 0$.

[7] This term was introduced by L. H. Rice in his paper entitled Couche ranks in the general matrix, Journal of Mathematics and Physics, vol. 7 (1928), pp. 94 ff.

An asterisk on T will indicate that the indices of T have been assigned fixed values (T^*). If every index of T is given the value 1 we write $T = 1$ (omitting the asterisk); similarly we write $T = 2$, etc. *A minor* of A is a matrix obtained from A by allowing the indices of A to vary over subsets of their ranges. *A* j-*way layer* of A is a minor of A obtained by fixing $(n - j)$ indices, where n is the class of A. *A* layer obtained by letting $T = T^*$ is called a T-layer.

A *cube* matrix is a matrix which has the same range on each index.

The usual notions concerning zero matrices, equality of matrices, etc. will apply here.

3. Compositions. Let two matrices A and B be displayed as (a_{T_1T}), (b_{TT_2}) respectively. The matrix $C = (a_{T_1T}\, b_{TT_2})$ is the *composite* of A and B on T, and is denoted by $A \mid T \mid B$. C is said to be obtained from A and B by composition on T. If T contains no indices, the composition is a *zero-composition*, and following the notation used by Weyl[8] in treating the two way case it is denoted by $A \times B$.

By a method analogous to that used by Hilbert,[9] the form

$$F = a_{i_1 \cdots i_n T}\, b_{T i_{n+1} \cdots i_{n+m}}\, x^1_{i_1} \cdots x^{n+m}_{i_{n+m}},\ T = j_1 \cdots j_l, \tag{1}$$

can be obtained from the forms

$$G = a_{i_1 \cdots i_n i_1 \cdots i_l}\, x^1_{i_1} \cdots x^n_{i_n}\, y^1_{i_1} \cdots y^l_{i_l},$$

$$H = b_{j_1 \cdots j_l i_{n+1} \cdots i_{n+m}}\, y^{l+1}_{j_1} \cdots y^{2\,l}_{j_l}\, x_{i_{n+1}} \cdots x_{i_{n+m}},$$

by applying the following differential operator

$$\sum_{y^1_{j_1} \cdots y^{2\,l}_{j_l}} \frac{\partial^{2\,l}}{\partial y^1_{j_1} \cdots \partial y^l_{j_l}\, \partial y^{l+1}_{j_1} \cdots \partial y^{2\,l}_{j_l}}$$

to GH. Several properties of compositions follow at once from the properties of differential operators. Now if the coefficients are sufficiently restricted F can be obtained by making the homogeneous transformation[10]

$$y^1_{j_1} \cdots y^l_{j_l} = b_{j_1 \cdots j_l i_{n+1} \cdots i_{n+m}}\, x^{n+1}_{i_{n+1}} \cdots x^{n+m}_{i_{n+m}}$$

on the variables of G.

If $A = (a_{i_1 \cdots i_n})$, $B = (b_{i_1 \cdots i_n})$, we define $A \pm B$ to be the matrix $(a_{i_1 \cdots i_n} \pm b_{i_1 \cdots i_n})$. If T_1, T_2 are disjointed, $A \mid T_1 \mid (B \mid T_2 \mid C) = (A \mid T_1 \mid B) \mid T_2 \mid C$. Evidently $A \mid T_1 \mid (B \pm C) = A \mid T_1 \mid B \pm A \mid T_1 \mid C$.

---

[8] H. Weyl, Gruppentheorie und Quantenmechanik, Leipzig, 1931, pp. 80 ff.

[9] D. Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Leipzig, 1912, pp. 128, 129.

[10] Cremona transformations are contained in the class of homogeneous transformations.

4. **Identity matrices.** If in the matrix $\delta = (\delta_{T_1T_2})$, $T_1 = T_2$, $\delta^*_{T_1}{}^*_{T_2} = 1$ if $T_1^* = T_2^*$, and $\delta^*_{T_1}{}^*_{T_2} = 0$ if $T_1^* \neq T_2^*$, then δ is said to be an identity matrix on T_1, T_2. Evidently if $(a_{T_2T_3}) = (a_{T_1T_3})$, $(\delta_{T_1T_2}) \mid T_2 \mid (a_{T_2T_3}) = (a_{T_1T_3})$.

If t is a number, $t \times \delta$ is a scalar matrix.

5. **2-way ranks.** In a matrix A the T-layers $A_1, A_2, \cdots, A_r$ are said to be linearly independent if and only if

$$c_f \times A_f = 0$$

implies that $c_f = 0$ for $f = 1, 2, \cdots, r$. The *T-rank* of A is the maximum number of linearly independent T-layers of A. Evidently the T^s-rank of A is the same as the T-rank of A where T^s is obtained from T by making a substitution S on the indices of T. Further the T_1-rank of A is the rank of the 2-way display $(a_{T_1T_2})$, and is hence called a 2-way rank (assuming that T_2 is not vacuous). If T_2 is vacuous, the T_1-rank of $A = (a_{T_1})$ is 1, or zero.

A matrix A is said to be *singular* on T if the T-rank of A is less than the order of T. If α and T are disjointed, the α-rank of A is said to be *invariant* under T-composition with C if the α-rank of A is equal to the α-rank of $A \mid T \mid C$.

6. **Properties of 2-way ranks under general compositions.** The reduction of the problem of properties of a set of compositions to the problem of properties of a single composition is accomplished by means of the lemma below.

Let matrices A_i possess determinants $\mid A_i \mid$ on T_{i1}, T_{i2} for $i = 1, 2, \cdots, n$. Let D represent the determinant $\mid A_1 \times A_2 \times \cdots \times A_n \mid$ whose rows are denoted by the partition $T_{11} \cdots T_{n1}$ and columns by $T_{12} \cdots T_{n2}$. Finally let Π_i denote the order of T_{i1}. With these notations we have the

LEMMA. *The determinant* $D = \mid A_1 \times A_2 \times \cdots \times A_n \mid = \mid A_1 \mid^{\Pi_2 \cdots \Pi_n} \mid A_2 \mid^{\Pi_1\Pi_3 \cdots \Pi_n} \cdots \mid A_n \mid^{\Pi_1 \cdots \Pi_{n-1}}$.

Weyl[11] has shown that

$$(A_1 \times \delta_1) \mid T_{12}T_{21} \mid (\delta_2 \times A_2) = A_1 \times A_2,$$

where δ_1, δ_2 are identity matrices on partitions T_{21}, T'_{21} and T_{12}, T'_{12} respectively. The matrix $A_1 \times \delta_1$ displayed with T_{11}, T'_{21} as the partition of its rows is a diagonal matrix with Π_2 minors equal to A_1 on the diagonal. Hence $\mid A_1 \times \delta_1 \mid = \mid A_1 \mid^{\Pi_2}$. Also $\mid A_2 \times \delta_2 \mid = \mid A_2 \mid^{\Pi_1}$, whence $\mid A_1 \times A_2 \mid = \mid A_1 \mid^{\Pi_2} \mid A_2 \mid^{\Pi_1}$. The Lemma follows by induction.

Since the matrices $B_i = (b_{T_{i1}T_{i2}})$ of composition on $T_{i1} = 1, 2, \cdots, n$ form a matrix $(e_{T_{11}T_{21} \cdots T_{n1}, T_{12}T_{22} \cdots T_{n2}}) = (b_{T_{11}T_{12}}) \times (b_{T_{21}T_{22}}) \times \cdots$ of composition on $T = T_{11} \cdots T_{n1}$, we have the

COROLLARY. *If a matrix* $A = (a_{T_1T_2})$ *possesses a determinant* $\mid a_{T_1T_2} \mid$, *the*

[11] H. Weyl, Ibid.

corresponding determinant of C obtained from A under composition on T_{i1} $(i = 1, \cdots, n)$ simply covering T_2 is equal to

$$\left| a_{T_1T_2} \right| \left| b_{T_{11}T_{12}} \right|^{\Pi_2 \cdots \Pi_n} \cdots \left| b_{T_{n1}T_{n2}} \right|^{\Pi_1 \cdots \Pi_{n-1}},$$

where $B_i = (b_{T_{i1}T_{i2}})$ are the matrices of composition. The B_i are assumed to be square as displayed.

By letting $B_{n-j+1}, \cdots, B_n$ be identity matrices on $T_{n-j+1,1}, T_{n-j+1,2}, \cdots, T_{n1}, T_{n2}$, it is seen that compositions on $T_{11}, T_{21}, \cdots, T_{n-j,1}$ forming a subset of T_2 are also treated by the corollary.

The following theorem is obvious.

THEOREM 1. *The T and σ-ranks of $A = (a_{T\sigma})$ are equal.*

We also have

THEOREM 2. *If T_1 and T_3 are disjointed and T_2 is a subset of T_3, the T_1-rank of a T_3-layer of A is less than or equal to the T_1-rank of a T_2-layer of A.*

The following theorem is a slight generalization of a theorem of Rice.[12]

THEOREM 3. *If $T = T_1 \cdots T_s$, the T-rank of A is less than or equal to the product of the T_i-ranks of A where $i = 1, \cdots, s$.*

By *elementary compositions* on A we shall mean elementary transformations[13] on the layers of A. By these compositions, if r_i is the T_i-rank of A, r_i of the T_i-layers of A can be reduced to zero layers, leaving by Theorem 6 the T, T_1, $\cdots$, T_s-ranks of A invariant. A is then reduced to a matrix in which there are at most $r_1 \cdots r_s$ non-zero intersections of the $T_1, \cdots, T_s$-layers. But these intersections are T-layers, whence the theorem is proved.

We have

THEOREM 4. *Let three matrices A, B, C be given by $(a_{T_1T_2T_3})$, $(b_{T_3T_4T_5})$, $A \mid T_3 \mid B$ respectively. Let r_i', r_i'', and r_i represent the T_i-ranks of A, B, and C respectively. Let further r_{25} be the T_2T_5-rank of C. Then $r_{25} \leq r_2' r_5''$.*

By theorem 3 $r_{25} \leq r_2 r_5$. Now $C = A \mid T_1T_3 \mid (\delta \times B)$, where $\delta = (\delta_{T_1T_1'})$ is an identity matrix on T_1, T_1'. Since C is the ordinary product of two 2-way matrices, $r_2 \leq r_2'$. Similarly $r_5 \leq r_5''$.

THEOREM 5. *Let A and B be defined as in Theorem 4. If B is non-singular on T_3, the T_2-rank of A is invariant under composition on T_3 with B.*

If $\delta = (\delta_{T_1T_1'})$, as above, by the Lemma $\delta \times B$ is non-singular on T_1T_3. Hence $r_2 = r_2'$.

THEOREM 6. *The $T_iT_j \cdots T_l$-rank $r_{ij\ldots l}$ of $A = (a_{T_1 \ldots T_s})$ is invariant under a set of compositions on $T_1, \cdots, T_s$ with matrices $C_1 = (c_{T_1,\sigma_1}), \cdots, C_s = (c_{T_s\sigma_s})$, where C_i is non-singular on T_i for $i = 1, 2, \cdots, s$.*

Let $E = (e_{\alpha,\beta})$ be defined to be the matrix

$$(e_{T_iT_j \ldots T_l,\sigma_i\sigma_j \cdots \sigma_l}) = (c_{T_i\sigma_i}\, c_{T_j\sigma_j} \cdots c_{T_l\sigma_l}),$$

and

$$F = (f_{T-\alpha,\sigma-\beta}) = (c_{T_1\sigma_1}\, c_{T_2\sigma_2} \cdots c_{T_{i-1}\sigma_{i-1}}\, c_{T_{i+1}\sigma_{i+1}} \cdots c_{T_{j-1}}\, c_{T_{j+1}} \cdots),$$

[12] L. H. Rice, Ibid.

[13] See Bocher's Introduction to Higher Algebra, p. 55.

where $T - \alpha$, α simply cover $T = T_1 \cdots T_s$, and $\sigma - \beta$, β simply cover $\sigma = \sigma_1 \cdots \sigma_s$. Now F is non-singular on $T - \alpha$ for we can choose subsets $\sigma_1', \sigma_2', \cdots$, of $\sigma_1, \sigma_2, \cdots$ respectively such that $(c_{T_1\sigma_1'})$ is non-singular on T_1, σ_1', $(c_{T_2\sigma_2'})$ is non-singular on T_2, σ_2', $\cdots$. By the Lemma the minor $(c_{T_1\sigma_1'}\ c_{T_2\sigma_2'} \cdots c_{T_{i-1}\sigma_{i-1}'}\ c_{T_{i+1}\sigma_{i+1}'} \cdots)$ of F is non-singular on $T - \alpha$, γ where $\gamma = \sigma_1'\sigma_2' \cdots \sigma_{i-1}'\sigma_{i+1}' \cdots$. By Theorem 5 $r_{ij \ldots l}$ is invariant under composition with F on $T - \alpha$. Similarly E is non-singular on α, whence we again apply Theorem 5.

The following Corollary to Theorem 6 was proved by Hitchcock[14] for polyadics.

COROLLARY. *The 2-way ranks of a matrix A are invariant under non-singular linear transformations on the indices of A.*

7. **Properties of ranks under zero-composition.** We are able to reduce the problem of ranks under zero-compositions to the problem under T-compositions where T is not vacuous. Let $C = A \times B$, where $C = (c_{\alpha\beta})$, $A = (a_\alpha)$, and $B = (b_\beta)$. Evidently $C = A' \mid i \mid B'$, where $A' = (a_{\alpha 1}')$, and $B' = (b_{\beta i}')$; furthermore i ranges over the single value 1, $a_{\alpha 1}' = a_\alpha$, and $b_{\beta 1}' = b_\beta$. The ranks of A and B are equal to the corresponding ranks of A' and B'.

Keeping the conventions of Theorem 4 concerning the r's we have

THEOREM 7. *If $A = (a_{T_1T_2})$, $B = (b_{T_3T_4})$, the T_2T_4-rank r_{24} of $A \times B$ satisfies the condition*

$$r_{24} \leqq r_2' r_4''.$$

The Theorem follows from Theorem 4 and the representations which involve C above.

COROLLARY 1. *If $A = (a_\alpha)$, $B = (b_\beta)$, the α- and β-ranks of the composite $A \times B$ are both unity or zero.*

For polyadics Corollary 1 is a consequence of a theorem of Hitchcock.[15]

By Theorem 4, $r_2 \leq r_2'$ where r_2, r_2' are the T_2-ranks of $A \times B$, A respectively. If B is not zero, some element $b_{T_3^* T_4^*} \neq 0$. Hence the T_2-rank of the layer $(a_{T_1T_2}\, b_{T_3^* T_4^*})$ is r_2'. We therefore have

COROLLARY 2. *If $A = (a_{T_1T_2})$, $B = (b_{T_3T_4})$ are not zero, the T_2-rank of $A \times B$ is equal to the T_2-rank of A.*

THEOREM 8. *If $A = (a_T)$ is not zero, and $B = (b_{\rho\sigma})$, the $T\sigma$-rank of $A \times B$ is equal to the σ-rank of B.*

Under composition on T with a matrix $C = (c_{T\alpha})$ non-singular on T,α the matrix $A \times B$ can be reduced to $A' \times B$ where A' contains an element unity while all other elements vanish. The $\alpha\sigma$-rank of $A' \times B$ is equal to the σ-rank of B. By Theorem 5 the $\alpha\sigma$-rank of $A' \times B$ is equal to the $T\sigma$-rank of $A \times B$, whence Theorem 8 follows.

---

[14] F. L. Hitchcock, The Expression of a Tensor or Polyadic as a sum of Products, Journal of Mathematics and Physics, vol. 6 (1927), p. 177.

[15] Hitchcock, Ibid.

8. **Inverse matrices.** Let $A = (a_{T_1T_2})$ be non-singular on T_1, T_2. The inverse of A on T_1, T_2 is defined to be the matrix $A^{-1}_{T_1T_2}$ which satisfies the relations

$$A \mid T_1 \mid A^{-1}_{T_1T_2} = \delta_{T_2}; A \mid T_2 \mid A^{-1}_{T_1T_2} = \delta_{T_1}, \tag{2}$$

where $\delta_{T_2} = (\delta_{T_2T_2'})$ is an identity matrix on T_2, T_2', and $\delta_{T_1} = (\delta_{T_1T_1'})$ is an identity matrix on T_1, T_1'.

THEOREM 9. *If a matrix $A = (a_{T_1T_2})$ is non-singular on T_1, T_2, there exists a unique inverse matrix*

$$A^{-1}_{T_1T_2} = (b_{T_2T_1}),$$

where $b_{T_2T_1}$ is the quotient of the cofactor of the element $a_{T_1T_2}$ in the 2-way display $(a_{T_1T_2})$ by the determinant of A on T_1, T_2.

9. **Matrices with ranks equal to unity.** The following theorem has been proved by Hitchcock[16] for polyadics.

THEOREM 10. *If the T-rank of a matrix $A = (a_{T\sigma})$ is unity, then A is the zero composite of two matrices one of whose partitions is the partition T.*

Evidently there exist matrices $B = (b_{\alpha T})$, $C = (c_{\sigma\beta})$, non-singular on α, T and σ, β respectively, such that

$$B \mid T \mid A \mid \sigma \mid C = A', \tag{3}$$

where $A' = (a'_{\alpha\beta})$, $a'_{\alpha^*\beta^*} = 1$ for $\alpha^* = \beta^* = 1$, while $a'_{\alpha^*\beta^*} = 0$ for $(\alpha^*, \beta^*) \neq (1,1)$. From (3) we have

$$A = B^{-1}_{\alpha T} \mid \alpha \mid A' \mid \beta \mid C^{-1}_{\sigma\beta}. \tag{4}$$

If we let $B^{-1}_{\alpha T} = (b^{\alpha T})$, $C^{-1}_{\sigma\beta} = (c^{\sigma\beta})$, then by (4) $A = (b^{1T}\, c^{\sigma 1})$, whence $A = (b^{1T}) \times (c^{\sigma 1})$. This proves the theorem.

If the matrix A of a multilinear form G satisfies the relation

$$A = B \times C,$$

it also satisfies the relation

$$A = kB \times (1/k)\, C,$$

where k is a scalar. Hence the form G factors uniquely only if a further restriction is placed upon G. We therefore have

THEOREM 11. *If the leading coefficient of a multilinear form G is unity, the multilinear form is factorable into 2 unique factors with leading coefficients unity if and only if the T-rank of G[17] is unity, where T is the total partition of the indices of one of the factors of G.*

---

[16] F. L. Hitchcock, Ibid.

[17] By ranks, partitions, ··· of G we mean the ranks, partitions, ··· of the matrix A associated with G.

The factorization property of Theorem 10 gives

THEOREM 12. *Let the partitions $\rho_1, \cdots, \rho_n$ of a matrix $A = (a_{T\sigma})$ of composition on T satisfy the relations $\rho_i = T_i\,\sigma_i$ for $i = 1, 2, \cdots, n$, where $T = T_1T_2 \cdots T_n$, $\sigma = \sigma_1\sigma_2 \cdots \sigma_n$. If the $\rho_1, \rho_2, \cdots, \rho_n$-ranks of A are unity, the composition with A on T is the same as a set of compositions $(b_{T_1\sigma_1})$, $(c_{T_1\sigma_2})$, $\cdots$, on the partitions $T_1, T_2, \cdots$, respectively.*

By Theorems 1 and 10 and Cor. 2 of Theorem 7 we have

THEOREM 13. *If the T-rank of a matrix $A = (a_{T\sigma})$ is unity, the α and β-ranks of A are equal, where α and β simply cover T.*

By Theorem 3 we have

THEOREM 14. *If the $T_1, \cdots, T_s$-ranks of a matrix $A = (a_{T_1 \cdots T_s\sigma})$ are all unity or zero, the T-rank of A is unity or zero, where $T_1, \cdots, T_s$ simply cover T.*

10. **Properties of identity matrices.** Identity matrices satisfy certain factorization properties. We have

THEOREM 15. *If $\delta = (\delta_{T_1T_2})$ is an identity matrix on T_1, T_2, where δ is of class $2r$, $T_1 = i_1 \cdots i_r$, $T_2 = j_1 \cdots j_r$, and if $\delta_1 = (\delta_{i_1j_1}), \cdots, \delta_r = (\delta_{i_rj_r})$ are identity matrices on $i_1, j_1, \cdots$, then δ is the zero-composite of r scalar matrices $s_1 \times \delta_1, \cdots, s_r \times \delta_r$, where $s_1 \cdots s_r = 1$.*

Evidently $\delta = \delta_1 \times \delta_2 \times \cdots \times \delta_r$. If we let $(a_{i_1j_1})$, $(a_{i_2j_2})$, $\cdots$, $(a_{i_rj_r})$ be the general factors of δ, we have

$$a_{i_1j_1}\, a_{i_2j_2} \cdots a_{i_rj_r} = \delta_{i_1 \cdots i_r\, j_1 \cdots j_r}. \tag{5}$$

Solution of (5) gives the scalar matrices of the theorem.

If from a partition T we delete all partitions which simply cover T, we shall say that we obtain a vacuous partition of order 1. We are now able to prove

THEOREM 16. *The T-rank of δ, where δ is an identity matrix on $i_1 \cdots i_n$, $j_1 \cdots j_n$, is equal to the order of the subset of T obtained by deleting all partitions of the type $i_t\, j_t$ (t not summed) from T.*

Consider the T-rank of $\delta = (\delta_{i_1i_2 \cdots i_n\, j_1j_2 \cdots j_n})$. Let the indices in T obtained after deleting all partitions of the type $i_t\, j_t$ be given by $i_u, i_w, \cdots, j_\alpha, j_\beta, \cdots$. Let the partition T' be obtained from T by replacing $j_\alpha, j_\beta, \cdots$ by $i_\alpha, i_\beta, \cdots$. By the symmetry of δ the T and T'-ranks of δ are equal. By Theorems 8 and 15 the T'-rank of δ is equal to the T_1^1-rank of δ', where δ' is an identity matrix on T_1^1, T_1^2; $T_1^1 = i_u\, i_w \cdots i_\alpha\, i_\beta \cdots$, $T_1^2 = j_u\, j_w \cdots j_\alpha\, j_\beta \cdots$. Let Π_1' be the order of T_1^1. The T_1^1-rank of δ' is Π_1', which proves the theorem.

11. **Ranks of inverse matrices.** From the results of §10 we shall prove

THEOREM 17. *If $A = (a_{T_1T_2})$ has an inverse on T_1, T_2, then A is non-singular on any partition T contained in T_1 or T_2.*

Let $\delta_1 = (\delta_{T_1T_1'})$ be an identity matrix on T_1, T_1'. By definition

$$A \mid T_2 \mid A^{-1}_{T_1T_2} = \delta_1.$$

By Theorem 5 the T-rank of δ_1 is equal to the T-rank of A, where T is contained in T_1. By Theorem 16 δ_1 is non-singular on T.

12. **Canonical multilinear forms.** By Theorem 6 and well-known theory on canonical bilinear forms we have

THEOREM 18. *Let a multilinear form F be given by*

$$a_{i_1 \cdots i_n} x^1_{i_1} \cdots , x^n_{i_n}.$$

Let the partitions $T_{\alpha\beta} = i_\alpha\, i_\beta$, $T_{\gamma\delta} = i_\gamma\, i_\delta$, $\cdots$, $T_{\xi\chi} = i_\xi\, i_\chi$ be mutually exclusive partitions of the indices of F. Let the i_ν-rank of F be denoted by r_ν for $\nu = 1, 2, \cdots, n$. Let $i_\rho, \cdots, i_\mu$ be such that $T_{\alpha\beta}, T_{\gamma\delta}, \cdots, T_{\xi\chi}, i_\rho, \cdots, i_\mu$ simply cover the total partition of the indices of F. If the $T_{\alpha\beta}, T_{\gamma\delta}, \cdots T_{\xi\chi}, i_\rho, \cdots, i_\mu$-ranks of F are unity, F is equivalent under non-singular linear transformations in the field of F to the form

$$K = b_{T_{\alpha\beta} T_{\gamma\delta} \cdots T_{\xi\chi}\, i_\rho \cdots i_\mu}\, y^\rho_{i\rho} \cdots y^\mu_{i\mu}\, w_{i\alpha}\, z_{i\beta}\, u_{i\gamma}\, v_{i\delta} \cdots x_{i\xi}\, t_{i\chi}, \tag{6}$$

where $b_{T^*_{\alpha\beta} \cdots T^*_{\xi\chi}\, i_\rho \cdots i_\mu} = 1$ for $i_\rho = \cdots = i_\mu = 1$; $i_\alpha = i_\beta = 1, 2, \cdots, r_\alpha$; $i_\gamma = i_\delta = 1, 2, \cdots, r_\gamma$; $\cdots$; $i_\xi = i_\chi = 1, 2, \cdots, r_\xi$; $b_{T^*_{\alpha\beta} \cdots T^*_{\xi\rho}\, i_\rho \cdots i_\mu} = 0$ for all others values of the subscripts.

If we take the matrix L of K to have the same range for each index as the corresponding index of F, then L is an identity matrix or an identity matrix bordered by zeros. Since K is determined by the $i_1, \cdots, i_n, T_{\alpha\beta}, T_{\gamma\delta}, \cdots, T_{\xi\chi}$-ranks of F, the remaining ranks of F are determined by these. Let σ be the partition obtained from the partition T of F by deleting all subsets of the type $T_{\alpha\beta}, \cdots, T_{\xi\chi}, i_\rho, \cdots, i_\mu$. It is readily shown that the T-rank of F is equal to the σ-rank of F. By Theorem 18 the σ-rank of F is equal to the product of the $i_j, i_f, \cdots, i_g$-ranks of F, where $i_j, i_f, \cdots, i_g$ simply cover σ.

When the matrix A of F is non-singular on all of its indices and $T_{\alpha\beta}, T_{\gamma\delta}, \cdots, T_{\xi\chi}$ simply cover the partition of all of the indices of F, then L is an identity matrix, and the matrices of transformation to identity form a matrix $A^{-1}_{T_1 T_2}$ such that

$$A \mid T_2 \mid A^{-1}_{T_1 T_2} = \delta;\ A \mid T_1 \mid A^{-1}_{T_1 T_2} = \delta, \tag{7}$$

where $L = \delta = (\delta_{T_1 T_2})$ is an identity matrix on T_1, T_2.

The canonical forms (6) are the only forms which are completely characterized by 2-way ranks as invariants for non-singular linear transformations.

13. **Note on the lemma.** By the Lemma and the equivalence of a set of linear transformations to a single composition we are able to extend the entire theory of the invariants of 2-way matrices with slight modifications to n-way matrices where $n \geq 2$. For example, from the theory of pairs of bilinear forms[18] we have at once.

[18] L. E. Dickson, Modern Algebraic Theories, p. 113 ff.

THEOREM 19. *Let 2 pairs of multilinear forms P_1, P_2 have the matrices $[A = (a_{\tau_1\tau_2}), B = (b_{\tau_1\tau_2})]$ and $[C = (c_{\tau_1\tau_2}), D = (d_{\tau_1\tau_2})]$ respectively. Let two 2-way matrices E, F be given by*

$$E = (e_{\tau_1\tau_2}) = (\rho a_{\tau_1\tau_2} + \sigma b_{\tau_1\tau_2})$$

$$F = (f_{\tau_1\tau_2}) = (\rho c_{\tau_1\tau_2} + \sigma d_{\tau_1\tau_2}).$$

If the determinants of E and F are not identically zero, and if the pairs of multilinear forms P_1, P_2 are equivalent under non-singular linear transformations, then the matrices E and F possess the same invariant factors.

Since n-way matrices, $n > 2$, may be displayed in 2-way form in several ways there may exist several sets of invariant factors for a pair of forms.

14. **Space determinants.** Hitchcock[19] has shown that certain space ranks of a polyadic P are invariant under non-singular linear transformations by exhibiting multiple concomitants which have as coefficients the minors of a space determinant associated with P. Using theorems on the space determinants of $A \mid T \mid B$, we shall prove relations between the space ranks of the matrices $A \mid T \mid B$, A, and B which include the ranks and the invariant properties established by Hitchcock.

We shall need a few notations. Let $(\alpha \cdots \gamma)$ be a substitution on the set of numbers $(1 \cdots n)$, where $\alpha, \cdots, \gamma = 1, \cdots, n$ in some order. Let $[\alpha \cdots \gamma]$ be $+$ or $-$ according as $(\alpha \cdots \gamma)$ is an even or odd substitution. Let $(\epsilon \cdots \mu)$, S be further substitutions on $(1 \cdots n)$. Let the product substitutions $S(\alpha \cdots \gamma)$, $S(\epsilon \cdots \mu)$ be represented by $(\alpha' \cdots \gamma')$, $(\epsilon' \cdots \mu')$ respectively. Evidently

$$[\alpha \cdots \gamma\,][\epsilon \cdots \mu] = [\alpha' \cdots \gamma'][\epsilon' \cdots \mu']. \tag{8}$$

If a p-way determinant D of order n is the expansion of $\mid a_{i_1 i_2 \cdots i_p} \mid$ with the indices $i_1, \cdots, i_r$ signant,[20] by the notation of (8)

$$D = \sum_{\substack{(i_{11} \cdots i_{1n}) \\ \vdots \\ (i_{p-1,1} \cdots i_{p-1,n})}} [i_{11} \cdots i_{1n}][i_{21} \cdots i_{2n}] \cdots [i_{r1} \cdots i_{rn}] \tag{9}$$

$$a_{i_{11} \cdots i_{p1}} \cdots a_{i_{1n} \cdots i_{pn}},$$

where the summation is over all substitutions on $(i_{11} \cdots i_{1n}), \cdots, (i_{p-1,1} \cdots i_{p-1,n})$; $(i_{11} \cdots i_{1n}), \cdots, (i_{p1} \cdots i_{pn})$ are substitutions on $(1 \cdots n)$. By (8) D is unique if r is even.[21] Hence we always take r even. The following theorem on extended Cayley products is proved by Rice.[22]

---

[19] F. L. Hitchcock, Multiple invariants and generalized rank of a p-way matrix of tensor, Journal of Mathematics and Physics, vol. 7 (1927), pp. 55–63.

[20] Defined by L. H. Rice, Introduction to Higher Determinants, Journal of Mathematics, and Physics, vol. 9 (1930), p. 48.

[21] Also mentioned by Rice, Ibid.

[22] L. H. Rice, Ibid.

THEOREM 20. *Let an nth-order determinant D be given by*

$$\left| \sum_{i_p=1}^{m} a_{i_1 \cdots i_p} b_{i_p j_2 \cdots j_q} \right|, m \geq n,$$

where an odd number of the i's and an odd number of the j's are signant. Let Γ *denote an nth-order subset of* $1, \cdots m$. *Let* A_Γ, B_Γ *denote nth-order minors of* $(a_{i_1 \cdots i_p})$, $(b_{i_p j_2 \cdots j_q})$ *obtained by letting* i_p *range over the subset* Γ. *Let* $| A_\Gamma |$, $| B_\Gamma |$ *denote the determinants of* A_Γ, B_Γ *expanded with* i_p *signant, and* $i_1, \cdots, i_{p-1}, j_2, \cdots, j_q$ *carrying the same signancy as in D. Then*

$$D = \sum_{\Gamma} | A_\Gamma | | B_\Gamma |,$$

where Γ *varies over the class of all nth order subsets of* $1, 2, \cdots, m$.

If $n > m$, D in Theorem 20 is zero. We shall also need

THEOREM 21. *Let D be an nth-order determinant given by*

$$\left| \sum_{i_p=1}^{m} a_{i_1 \cdots i_p} b_{i_p j_2 \cdots j_q} \right|,$$

where an even number of the i's and an even number of the j's are signant. Let A_Γ *denote an nth-order matrix* $(a_{i_1 \cdots i_{p-1} \rho})$ *determined by letting* ρ *vary over the range* Γ, *where* Γ *is an nth-order subset of* $1, \cdots, m$, *repetitions of values allowed. Let* $| A_\Gamma |$ *be the determinant of* A_Γ *expanded with* ρ *non-signant. Similarly, let* $| B_\Gamma |$ *be the determinant of an nth-order matrix* $(b_{\rho j_2 \cdots j_q})$ *expanded with* ρ *non-signant. The remaining indices of* $| A_\Gamma |$ *and* $| B_\Gamma |$ *carry the same signancy as in D. Then*

$$D = \sum_{\Gamma} 1/k_\Gamma | A_\Gamma | | B_\Gamma |,$$

where Γ *varies over all nth-order subsets of* $1, 2, \cdots, m$, *repetitions of values allowed, and the numbers of* k_Γ *are integers.*

It is no restriction to assume that $i_1, i_2, \cdots, i_s, j_2, j_3, \cdots, j_t$ are signant in D while the remaining indices are non-signant. By (9) D can be written as

$$(10) \quad \sum_{i_{p1} \cdots i_{pn}} \left(\sum_{\substack{(i_{11} \cdots i_{1n}) \\ \vdots \\ (i_{p-1,1} \cdots i_{p-1,n})}} [i_{11} \cdots i_{1n}] \cdots [i_{s1} \cdots i_{sn}] a_{i_{j1} \cdots i_{p1}} \cdots a_{i_{1n} \cdots i_{pn}} \right)$$

$$\left(\sum_{\substack{(j_{21} \cdots j_{2n}) \\ \vdots \\ (j_{q-1,1} \cdots j_{q-1,n})}} [j_{21} \cdots j_{2n}] \cdots [j_{t1} \cdots j_{tn}] b_{i_{p1} j_{21} \cdots i_{q1}} \cdots b_{i_{pn} j_{2n} \cdots i_{qn}} \right).$$

Of the factors exhibited in (10), those obtained by letting $i_{p1}, i_{p2}, \cdots, i_{pn}$ have distinct values are determinants $| A_\Gamma |$, $| B_\Gamma |$. The terms in (10) which are

such that $i_{p1} = i_{p2} = \cdots = i_{pr}$, $i_{p,r+1} = i_{p,r+2} = \cdots = i_{p,r+u}$, $\cdots$ can be written $1/k_\Gamma \mid A_\Gamma \mid\mid B_\Gamma \mid$, where $k_\Gamma = r!\,u!\cdots w!$ This proves the theorem.

15. **Properties of space ranks under general compositions.** Let a matrix A be given by $(a_{\tau_1 \cdots \tau_\rho \tau_{\rho+1} \cdots \tau_p})$. Let the partition $T_{\rho+t}$ of A be given by the set of indices $i_{t1} \cdots i_{ts_t}$ for $t = 1, \cdots, p-\rho$. Let Π be the smallest of the orders of $T_1, \cdots, T_\rho$, and let n_{tq} be the order of i_{tq} in $T_{\rho+t}$ for $q = 1, \cdots, s_t$, $t = 1, \cdots, p-\rho$. Let partitions $\alpha_{\rho+t}$ be given by $\alpha_{\rho+t} = j_{t1} \cdots j_{ts_t}$, $t = 1, \cdots, p-\rho$, where the index j_{tq} ranges over $1, \cdots, n_{tq}, n_{tq}+1, \cdots 2n_{tq}, \cdots, \Pi n_{tq}$. Let further a matrix A' be given by $(a'_{\tau_1 \cdots \tau_\rho \alpha_{\rho+1} \cdots \alpha_p})$ where $a'_{\tau_1^* \cdots \tau_\rho^* \alpha_{\rho+1}^* \cdots \alpha_p^*} = a_{\tau_1^* \cdots \tau_\rho^* \tau_{\rho+1}^* \cdots \tau_p^*}$ when every index $j_{tq} \equiv i_{tq} \pmod{n_{tq}}$ for $q = 1, \cdots, s_t$, $t = 1, \cdots, p-\rho$. The matrix A' will be called the $(T_1 \cdots T_\rho, T_{\rho+1} \cdots T_p)$-*derivate* of A.

Examples. In the case of a rectangular two-way matrix $B = (b_{ij})$ of order n by m the $(ij,)$-derivate of B is B itself; the (i, j)-derivate of B (which cannot be used in the treatment of ranks below) is given by

$$(B_1 B_2 \cdots B_n) = \begin{bmatrix} b_{11} \cdots b_{1m} & b_{11} \cdots b_{1m} \cdots b_{1m} \\ \cdot & \\ \cdot & \\ \cdot & \\ b_{n1} \cdots b_{nm} & b_{11} \cdots b_{nm} \cdots b_{nm} \end{bmatrix},$$

where $B_1 = B_2 \cdots = B_n = B$.

We shall need the

DEFINITION. The p-way *rank*[23] $r(T_1 \cdots T_\rho, T_{\rho+1} \cdots T_p)$, $\rho \geq 2$, of $A = (a_{\tau_1 \cdots \tau_p})$ is the order of the largest non-vanishing minor of the $(T_1 T_2 \cdots T_\rho, T_{\rho+1} \cdots T_p)$-derivate of A, where $T_1, T_2, \cdots, T_\rho$ are signant and $\alpha_{\rho+1}, \alpha_{\rho+2}, \cdots, \alpha_p$ are non-signant.

If all partitions are signant we have $r(T_1 \cdots T_p)$.

Let $T_1^{S_1}, \cdots, T_p^{S_p}$ be partitions obtained from $T_1, \cdots, T_p$ by making the substitutions $S_1, \cdots, S_p$ on the indices of $T_1, \cdots, T_p$ respectively. Evidently the rank $r(T_1^{S_1} \cdots T_\rho^{S_\rho}, T_{\rho+1}^{S_{\rho+1}} \cdots T_p^{S_p})$ is equal to the rank $r(T_1 \cdots T_\rho, T_{\rho+1} \cdots T_p)$ of A.

We shall now prove

THEOREM 22. *Let three matrices A, B, C be given by $(a_{\tau_1 \cdots \tau_\sigma})$, $(b_{\tau_\sigma \gamma_1 \cdots \gamma_\omega})$, $A \mid T_\sigma \mid B$, respectively. Let $\delta_1 = T_{\epsilon+1}\,\gamma_{\lambda+1}, \cdots, \delta_{\nu+\chi} = T_{\epsilon+\nu+\chi}\,\gamma_{\lambda+\nu+\chi}$. Let the matrix $\delta = (\delta_{T_{\epsilon+1} T'_{\epsilon+1}}) \times \cdots \times (\delta_{T_{\epsilon+\nu+\chi} T'_{\epsilon+\nu+\chi}})$, where $(\delta_{T_{\epsilon+s} T'_{\epsilon+s}})$ is an identity matrix on $T_{\epsilon+s}, T'_{\epsilon+s}$ for $s = 1, \cdots, \nu+\chi$. Let partitions $\delta'_1 = T'_{\epsilon+1}\,\gamma_{\lambda+1}, \cdots, \delta'_{\nu+\chi} = T'_{\epsilon+\nu+\chi}\,\gamma_{\lambda+\nu+\chi}$. Let further $\rho = T_{\epsilon+1} \cdots T_{\epsilon+\nu+\chi}\,T_\sigma$. Then*

1. *if ϵ, $\nu + \lambda$ are both odd the rank $r(\delta_1 \cdots \delta_\nu\, T_1 \cdots T_\epsilon\, \gamma_1 \cdots \gamma_\lambda, T_{\epsilon+\nu+\chi+1} \cdots T_{\sigma-1}\,\gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta_{\nu+1} \cdots \delta_{\nu+\chi})$ of C is less than or equal to the rank $r(T_1 \cdots$*

[23] The ranks defined by Hitchcock, Op. Cit., are the special cases of $r(T_1 \cdots T_\rho, T_{\rho+1} \cdots T_p)$ where $\rho = p$ or $p-1$.

$T_\epsilon\ \rho,\ T_{\epsilon+\nu+\chi+1} \cdots T_{\sigma-1}$) *of* A *and the rank* $r(\delta_1' \cdots \delta_\nu'\ \gamma_1 \cdots \gamma_\lambda\ \rho,\ \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\ \delta_{\nu+1}' \cdots \delta_{\nu+\chi}')$ *of the matrix* $B \times \delta$;

2. *if* $\epsilon,\ \nu + \lambda$ *are both even, and* $\epsilon \geq 2,\ \nu + \lambda \geq 2$, *the rank* $r(\delta_1 \cdots \delta_\nu\ T_1 \cdots T_\epsilon\ \gamma_1 \cdots \gamma_\lambda,\ T_{\epsilon+\nu+\chi+1} \cdots T_{\sigma-1}\ \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\ \delta_{\nu+1} \cdots \delta_{\nu+\chi})$ *of* C *is less than or equal to the rank* $r(T_1 \cdots T_\epsilon,\ \rho T_{\epsilon+\nu+\chi+1} \cdots T_{\sigma-1})$ *of* A *and the rank* $r(\delta_1' \cdots \delta_\nu'\ \gamma_1 \cdots \gamma_\lambda,\ \rho\gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\ \delta_{\nu+1}' \cdots \delta_{\nu+\chi}')$ *of* $B \times \delta$.

The rank $r(\delta_1 \cdots \delta_\nu\ T_1 \cdots T_\epsilon \gamma_1 \cdots \gamma_\lambda,\ T_{\epsilon+\nu+\chi+1} \cdots T_{\sigma-1}\ \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\ \delta_{\nu+1} \cdots \delta_{\nu+\chi})$ of C is evidently the same as the rank $r(\delta_1' \cdots \delta_\nu'\ T_1 \cdots T_\epsilon \gamma_1 \cdots \gamma_\lambda,\ T_{\epsilon+\nu+\chi+1} \cdots T_{\sigma-1}\ \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\ \delta_{\nu+1}' \cdots \delta_{\nu+\chi}')$ of C_1, where

$$C_1 = A \mid \rho \mid B \times \delta.$$

Let $\mid M \mid_q$ be a qth-order minor of the $(\delta_1' \cdots \delta_\nu'\ T_1 \cdots T_\epsilon\ \gamma_1 \cdots \gamma_\lambda,\ T_{\epsilon+\nu+\chi+1} \cdots T_{\sigma-1}\ \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\ \delta_{\nu+1}' \cdots \delta_{\nu+\chi}')$-derivate of C_1 expanded with $\delta_1', \cdots, \delta_\nu', T_1, \cdots, T_\epsilon, \gamma_1, \cdots, \gamma_\lambda$ signant. If $\epsilon,\ \nu + \lambda$ are odd, we apply Theorem 20 to $\mid M \mid_q$; if $\epsilon,\ \nu + \lambda$ are even, we apply Theorem 21, whence Theorem 22 is proved.

The ordering of the signant and non-signant partitions used in Theorem 22 is no restriction of the problem since the partitions can be chosen at random from the indices of A. The ranks of C covered by this theorem compose the class of all possible ranks of C for which $\epsilon,\ \nu + \lambda \geq 1$.

In our discussion above we defined the p-way rank of a matrix A in terms of the minors of a derivate of A where these minors possessed the property that at least two partitions were signant. By constructing new derivates of A we shall now define p-way ranks of A in terms of minors which have no partitions signant. These ranks will be called *non-signant ranks*. By means of these ranks we shall be able to treat the ranks of C occurring in Theorem 22 for which $\epsilon = 0$ or $\nu + \lambda = 0$.

Let a matrix A be given by $(a_{T_1 \ldots T_p})$ as before. Let $\mu_1, \cdots, \mu_p$ be arbitrary positive integers. Let the partition T_t of A be given by the set of indices $i_{t1} \cdots i_{ts_t}$ for $t = 1, \cdots, p$, and let n_{tq} be the order of i_{tq} for $q = 1, \cdots, s_t$, $t = 1, \cdots, p$. Let partitions α_t be given by $\alpha_t = j_{t1} \cdots j_{ts_t}$ for $t = 1, \cdots, p$. Let a matrix $A' = (a'_{\alpha_1 \cdots \alpha_p})$ satisfy the following properties;

1. The index j_{tq} ranges over $1, \cdots, n_{tq},\ n_{tq} + 1, \cdots, 2n_{tq}, \cdots \mu_t n_{tq}$ for $q = 1, \cdots, s_t, t = 1, \cdots, p$;

2. $a'_{\alpha_1^* \ldots \alpha_p^*} = a_{T_1^* \ldots T_p^*}$ when every index $j_{tq} \equiv i_{tq} \pmod{n_{tq}}$ for $q = 1, \cdots, s_t,\ t = 1, \cdots, p$.

The matrix A' so defined is called the $(,T_1 \cdots T_p; \mu_1 \cdots \mu_p)$-derivate of A.

Example. The $(,ij; nm)$-derivate of $B = (b_{ij})$ is given by

$$\begin{bmatrix} B_{11} & \cdots & B_{1m} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ B_{n1} & \cdots & B_{nm} \end{bmatrix},$$

where $B_{\rho\sigma} = B$ for $\rho = 1, \cdots, n; \sigma = 1, \cdots, m$.

Let a minor $| M |$ of the $(,T_1 \cdots T_p; \mu_1 \cdots \mu_p)$-derivate of $A = (a_{T_1 \cdots T_p})$ satisfy the properties:

1. It is of order less than or equal to α.

2. It has no partitions signant.

DEFINITION. The p-way rank $r(,T_1 \cdots T_p; \mu_1 \cdots \mu_p; \alpha)$ of $A = (a_{T_1 \cdots T_p})$ is the order of the largest non-vanishing minor $| M |$ satisfying the properties 1 and 2.

Let $T_1^{s_1}, \cdots, T_p^{s_p}$ be defined as above. In the case of non-signant ranks it is again evident the rank $r(T_1^{s_1} \cdots T_p^{s_p}; \mu_1 \cdots \mu_p; \alpha)$ of A is equal to the rank $(,T_1 \cdots T_p; \mu_1 \cdots \mu_p; \alpha)$ of A.

In the remainder of the paper a minor $| M |_q$ of order q of the $(T_1 \cdots T_\rho, T_{\rho+1} \cdots T_p)$-derivate of a matrix $A = (a_{T_1 \cdots T_p})$ will denote a minor of this derivate obtained by letting $T_1, \cdots T_\rho$ be signant partitions, and $T_{\rho+1}, \cdots, T_p$ non-signant partitions in the expansion of this minor. It will also be understood that a minor $| M |_q$ of order q of the $(,T_1 \cdots T_p; \mu_1 \cdots \mu_p)$-derivate of A will denote a minor of this derivate obtained by letting all partitions be non-signant.

In the case where $p = 1$, so that $r(,T_1 \cdots T_p; \mu_1 \cdots \mu_p; \alpha)$ of A contains only one partition we have the evident

THEOREM 23. *If at least one element of $A = (a_T)$ is not zero, and if $\alpha \geq \mu$, the rank $r(,T; \mu; \alpha)$ of A is greater than or equal to μ. If only one element is not zero and $\alpha \geq \mu$, $r(,T; \mu; \alpha)$ of A is μ. If $\mu \geq \alpha$, in either case $r(,T; \mu; \alpha)$ of A is α.*

We shall now prove

THEOREM 24. *Let three matrices A, B, C be given by $(a_{T_1 \cdots T_\sigma})$, $(b_{T_{\sigma\gamma 1} \cdots \gamma_\omega})$, $A \mid T_\sigma \mid B$, respectively. Let $\delta_1 = T_{\epsilon+1}\, \gamma_1, \cdots, \delta_\chi = T_{\epsilon+\chi}\, \gamma_\chi$. Let the matrix $\delta = (\delta_{T_{\epsilon+1} T'_{\epsilon+1}}) \times \cdots \times (\delta_{T_{\epsilon+\chi} T'_{\epsilon+\chi}})$, where $(\delta_{T_{\epsilon+s} T'_{\epsilon+s}})$ is an identity matrix on $T_{\epsilon+s}, T'_{\epsilon+s}$, for $s = 1, \cdots, \chi$. Let $\delta'_1 = T'_{\epsilon+1}\, \gamma_1, \cdots, \delta'_\chi = T'_{\epsilon+\chi} \gamma_\chi$. Let further $\rho = T_{\epsilon+1} \cdots T_{\epsilon+\chi}\, T_\sigma$. Then the rank $r(T_1 \cdots T_\epsilon, T_{\epsilon+\chi+1} \cdots T_{\sigma-1}\, \gamma_{\chi+1} \cdots \gamma_\omega\, \delta_1 \cdots \delta_\chi)$, $\epsilon \geq 2$, of C is less than or equal to the rank $r(T_1 \cdots T_\epsilon, T_{\epsilon+\chi+1} \cdots T_{\sigma-1}\, \rho)$ of A and $r(,\rho\gamma_{\chi+1} \cdots \gamma_\omega\, \delta'_1 \cdots \delta'_\chi; \mu \cdots \mu; \mu)$ of $B \times \delta$, where μ is the smallest of the orders of $T_1, \cdots, T_\epsilon$.*

The rank $r(T_1 \cdots T_\epsilon, T_{\epsilon+\chi+1} \cdots T_{\sigma-1}\, \gamma_{\chi+1} \cdots \gamma_\omega\, \delta_1 \cdots \delta_\chi)$ of C is equal to the rank $r(T_1 \cdots T_\epsilon, T_{\epsilon+\chi+1} \cdots T_{\sigma-1}\, \gamma_{\chi+1} \cdots \gamma_\omega\, \delta'_1 \cdots \delta'_\chi)$ of C_1 where

$$C_1 = A \mid \rho \mid B \times \delta.$$

We now apply Theorem 21 to the minors of the $(T_1 \cdots T_\epsilon, T_{\epsilon+\chi+1} \cdots T_{\sigma-1}\, \gamma_{\chi+1} \cdots \gamma_\omega\, \delta'_1 \cdots \delta'_\chi)$-derivate of C_1.

We have at once the

COROLLARY. *Let matrices A, B, C be given as in Theorem 24. Let $\delta_1, \cdots, \delta_{\nu+\chi}, \delta'_1, \cdots, \delta'_{\nu+\chi}, \delta, \rho$ be given as in Theorem 22 except that $\epsilon = 0$. Then the rank $r(\delta_1 \cdots \delta_\nu\, \gamma_1 \cdots \gamma_\lambda, T_{\nu+\chi+1} \cdots T_{\sigma-1}\, \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta_{\nu+1} \cdots \delta_{\nu+\chi})$, $\nu + \lambda > 2$, of C is less than or equal to the rank $r(\delta'_1 \cdots \delta'_\nu\, \gamma_1 \cdots \gamma_\lambda, \rho\gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$ of $B \times \delta$ and the rank $r(,\rho T_{\nu+\chi+1} \cdots T_{\sigma-1}; \mu\mu \cdots \mu; \mu)$ of A where μ is the smallest of the orders of $\delta_1, \cdots, \delta_\nu, \gamma_1, \cdots \gamma_\lambda$.*

For non-signant ranks we have

THEOREM 25. *Let three matrices A, B, C be given by $(a_{T_1 \cdots T_\sigma})$, $(b_{T_\sigma \gamma_1 \cdots \gamma_\omega})$, $A \mid T_\sigma \mid B$, respectively. Let $\delta_1 = T_{\epsilon+1}\,\gamma_{\lambda+1}, \cdots, \delta_\nu = T_{\epsilon+\nu}\,\gamma_{\lambda+\nu}$. Let the matrix $\delta = (\delta_{T_{\epsilon+1}\,T'_{\epsilon+1}}) \times \cdots \times (\delta_{T_{\epsilon+\nu}\,T'_{\epsilon+\nu}})$ where $(\delta_{T_{\epsilon+s}\,T'_{\epsilon+s}})$ is an identity matrix on $T_{\epsilon+s}$, $T'_{\epsilon+s}$ for $s = 1, \cdots, \nu$. Further let $\delta'_1 = T'_{\epsilon+1}\,\gamma_{\lambda+1}, \cdots, \delta'_1 = T'_{\epsilon+\nu}\,\gamma_{\lambda+\nu}$, and let $\rho = T_{\epsilon+1} \cdots T_{\epsilon+\nu}\,T_\sigma$. Then the rank $r(,\delta_1 \cdots \delta_\nu\, T_1 \cdots T_\epsilon\, \gamma_1 \cdots \gamma_\lambda;\ \mu_1 \cdots \mu_\nu\, \mu_{\nu+1} \cdots \mu_{\nu+\epsilon}\, \mu_{\nu+\epsilon+1} \cdots \mu_{\nu+\epsilon+\lambda};\ \alpha)$ of C is less than or equal to the rank $r(,\rho\delta'_1 \cdots \delta'_\nu\, \gamma_1 \cdots \gamma_\lambda;\ \alpha\mu_1 \cdots \mu_\nu\, \mu_{\nu+\epsilon+1} \cdots \mu_{\nu+\epsilon+\lambda};\ \alpha)$ of $B \times \delta$ and the rank $r(,\rho T_1 \cdots T_\epsilon;\ \alpha\mu_{\nu+1} \cdots \mu_{\nu+\epsilon};\ \alpha)$ of A.*

To prove Theorem 25 we construct a matrix $C_1 = A \mid \rho \mid B \times \delta$, and apply Theorem 21.

16. **The ranks of $B \times \delta$.** The ranks of the matrix $B \times \delta = (b_{T_\sigma \gamma_1 \cdots \gamma_\omega}\, \delta_{T_{\epsilon+1} \cdots T'_{\epsilon+1}} \cdots \delta_{T_{\epsilon+\nu+\chi}\,T'_{\epsilon+\nu+\chi}})$, $\nu + \chi > 0$, mentioned in Theorems 22 to 25 are not directly expressible for a general B in terms of the ranks of B. We shall, however, establish a bound property between the ranks of these matrices. We consider several cases.

1. ρ signant. Assume that a qth-order minor $\mid B \mid_q$ of the $(T_\sigma\, \gamma_1 \cdots \gamma_{\lambda+\nu}, \gamma_{\lambda+\nu+1} \cdots \gamma_\omega)$-derivate of B is not zero. Let a minor $\mid B \times \delta \mid_q$ of order q be obtained from the $(\delta'_1 \cdots \delta'_\nu\, \gamma_1 \cdots \gamma_\lambda \rho,\ \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$-derivate of $B \times \delta$ by letting $T_\sigma, \gamma_1, \cdots, \gamma_\omega$ have the same ranges as in $\mid B \mid_q$, and choose the ranges of $\delta'_1, \cdots, \delta'_{\nu+\chi}$, such that $T_{\epsilon+i} = T'_{\epsilon+i} = 1$ for $i = 1, \cdots, \nu + \chi$. Evidently

$$\mid B \times \delta \mid_q = \mid B \mid_q.$$

Hence the rank $r(\delta'_1 \cdots \delta'_\nu\, \gamma_1 \cdots \gamma_\lambda\, \rho,\ \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$ of $B \times \delta$ is greater than or equal to the rank $r(T_\sigma\, \gamma_1 \cdots \gamma_{\lambda+\nu}, \gamma_{\lambda+\nu+1} \cdots \gamma_\omega)$ of B.

1_1. In special cases a better inequality than that of Case 1 can be obtained as follows. Let μ be the smallest of the orders of $T_{\epsilon+1}, \cdots, T_{\epsilon+\nu+\chi}$. Assume that $\lambda = 0$ in Case 1 and $\omega = \nu + \chi$, and further that a qth-order minor $\mid B \mid_q$ of the $(T_\sigma\, \gamma_1 \cdots \gamma_\nu, \gamma_{\nu+1} \cdots \gamma_{\nu+\chi})$-derivate of B is not zero. Let a minor $\mid B \times \delta \mid_{\mu q}$ of order μq be obtained from the $(\delta'_1 \cdots \delta'_\nu\, \rho, \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$-derivate of B by letting $T'_{\epsilon+i} = 1, 2, \cdots, \mu$ for $i = 1, 2, \cdots, \nu + \chi$; $(T_{\epsilon+1} \cdots T_{\epsilon+\nu+\chi}) = (1 \cdots 1), (2 \cdots 2), \cdots, (\mu \cdots \mu)$; and $T_\sigma, \gamma_1, \cdots, \gamma_{\nu+\chi}$ have the same ranges as in $\mid B \mid_q$. Then

$$\mid B \times \delta \mid_{\mu q} = \mid B \mid_q^\mu,$$

whence the rank $r(\delta'_1 \cdots \delta'_\nu\, \rho,\ \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$ of $B \times \delta$ is greater than or equal μr_1, where $r_1 = r(T_\sigma\, \gamma_1 \cdots \gamma_\nu, \gamma_{\nu+1} \cdots \gamma_{\nu+\chi})$ of B.

2. ρ non-signant; $\nu + \lambda \geq 2$. Applying the same argument as in Case 1 using now the $(\gamma_1 \cdots \gamma_{\lambda+\nu},\ T_\sigma\, \gamma_{\lambda+\nu+1} \cdots \gamma_\omega)$-derivate of B, and the $(\delta'_1 \cdots \delta'_\nu\, \gamma_1 \cdots \gamma_\lambda,\ \rho\gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$-derivate of $B \times \delta$, we can show that the rank $r(\delta'_1 \cdots \delta'_\nu\, \gamma_1 \cdots \gamma_\lambda,\ \rho\gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$ of $B \times \delta$ is greater than or equal to the rank $r(\gamma_1 \cdots \gamma_{\lambda+\nu}, T_\sigma\, \gamma_{\lambda+\nu+1} \cdots \gamma_\omega)$ of B.

2_1. By methods similar to Case 1_1 we can show that the rank $r(\delta_1' \cdots \delta_\nu', \rho\delta_{\nu+1}' \cdots \delta_{\nu+\chi}')$ of $B \times \delta$ is greater than or equal to μr_1, where $r_1 = r(\gamma_1 \cdots \gamma_\nu, T_\sigma \gamma_{\nu+1} \cdots \gamma_{\nu+\chi})$ of B, and μ is the smallest of the orders of $T_{\epsilon+1}, \cdots, T_{\epsilon+\nu+\chi}$.

3. No partitions signant. Let $\delta_1' = T_1' \gamma_1, \cdots, \delta_\chi' = T_\chi' \gamma_\chi$, $\rho = T_1 \cdots T_\chi T_\sigma$. Assume that $| B |_q \neq 0$, where $| B |_q$ is a minor of order q of the $(,T_\sigma \gamma_{\chi+1} \cdots \gamma_\omega \gamma_1 \cdots \gamma_\chi; \mu\mu_{\chi+1} \cdots \mu_\omega \mu_1 \cdots \mu_\chi)$-derivate of B. Let a minor $| B \times \delta |_q$ of order q of the $(,\rho\gamma_{\chi+1} \cdots \gamma_\omega \delta_1' \cdots \delta_\chi'; \mu\mu_{\chi+1} \cdots \mu_\omega \mu_1 \cdots \mu_\chi)$-derivate of $B \times \delta$ be obtained by letting $T_\sigma, \gamma_1, \cdots, \gamma_\omega$ have the same ranges as in $| B |_q$, and $T_{\epsilon+i}' = T_{\epsilon+i} = 1$ for $i = 1, \cdots, \chi$. Then

$$| B \times \delta |_q = | B |_q.$$

Hence the rank $r(,\rho\gamma_{\chi+1} \cdots \gamma_\omega \delta_1' \cdots \delta_\chi'; \mu\mu_{\chi+1} \cdots \mu_\omega \mu_1 \cdots \mu_\chi; \alpha)$ of $B \times \delta$ is greater than or equal to the rank $r(,T_\sigma \gamma_{\chi+1} \cdots \gamma_\omega \gamma_1 \cdots \gamma_\chi; \mu\mu_{\chi+1} \cdots \mu_\omega\mu_1 \cdots \mu_\chi; \alpha)$ of B.

3_1. A better inequality for special cases of 3 can be obtained as follows. Assume that $| B |_q \neq 0$, where $| B |_q$ is a minor of order q of the $(,T_\sigma\gamma_1 \cdots \gamma_\chi; \mu\mu_1 \cdots \mu_\chi)$-derivate of B. Let a minor $| B \times \delta |_{\Pi q}$ of order Πq of the $(,\rho\delta_1' \cdots \delta_\chi'; \Pi\mu, \Pi\mu_1, \cdots \Pi\mu_\chi)$-derivate of $B \times \delta$ be obtained by letting $T_\sigma, \gamma_1, \cdots, \gamma_\chi$ have the same ranges as in $| B |_q$; also $T_i' = 1, \cdots, \Pi$ for $i = 1, \cdots, \chi$, and $(T_1 \cdots T_\chi) = (1 \cdots 1), \cdots, (\Pi \cdots \Pi)$. Π is taken as an arbitrary integer greater than zero and less than the smallest of the orders of $T_1, \cdots, T_\chi$. Then

$$| B \times \delta |_{\Pi q} = | B |_q^{\Pi}.$$

Therefore the rank $r(,\rho\delta_1' \cdots \delta_\chi'; \Pi\mu, \Pi\mu_1, \cdots, \Pi\mu_\chi; \Pi\alpha)$ of $B \times \delta$ is greater than or equal to Πr_1, where $r_1 = r(,\rho\gamma_1 \cdots \gamma_\chi; \mu\mu_1 \cdots \mu_\chi, \alpha)$ of B.

This completes the treatment of the ranks of $B \times \delta$ mentioned in Theorems 22 to 25.

17. **Space ranks of $A \times B$.** The following theorem is given by Hitchcock.[24]

THEOREM 26. Let $A = (a_{i_1 \cdots i_p})$, $B = (b_{j_1 \cdots j_q})$, and $C = A \times B$, where A and B are cube matrices. If $| C |$ is the determinant of C developed with an even number of i's and j's signant, then

$$| C | = n! | A | | B |,$$

where $| A |$ and $| B |$ are determinants of A and B with the same signancy as in C.

Theorem 26 implies

THEOREM 27. *Let $A = (a_{\tau_1 \cdots \tau_p})$, $B = (b_{\gamma_1 \cdots \gamma_\omega})$, $C = A \times B$. If s, t are both even and greater than or equal to 2, the rank $r(T_1 \cdots T_s \gamma_1 \cdots \gamma_t, T_{s+1} \cdots T_p \gamma_{t+1} \cdots \gamma_\omega)$ of C is equal to the smaller of the ranks $r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$ and $r(\gamma_1 \cdots \gamma_t, \gamma_{t+1} \cdots \gamma_\omega)$ of A and B respectively.*

If s and t are both odd the rank $r(T_1 \cdots T_s \gamma_1 \cdots \gamma_t, T_{s+1} \cdots T_p \gamma_{t+1} \cdots \gamma_\omega)$

[24] F. L. Hitchcock, Ordered determinants with applications to polyadics, Journal of Mathematics and Physics, vol. 4 (1925), p. 213.

of C is 1 or 0 according as A and B contain non-zero elements or at least one of these matrices is a zero matrix.

THEOREM 28. *Let* $A = (a_{\tau_1 \cdots \tau_p})$, $B = (b_{\gamma_1 \cdots \gamma_\omega})$, *and* $C = A \times B$. *Let* Π *be the smallest of the orders of the partitions* $T_1, \cdots, T_s$. *The rank* $r(T_1 \cdots T_s, T_{s+1} \cdots T_p \gamma_1 \cdots \gamma_\omega)$ *of* C *is equal to the smaller of the ranks* $r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$ *of* A *and* $r(,\gamma_1 \cdots \gamma_\omega; \Pi \cdots \Pi; \Pi)$ *of* B.

Theorem 28 and the following theorem follow from Theorem 26.

THEOREM 29. *Let* $A = (a_{\tau_1 \cdots \tau_p})$, $B = (b_{\gamma_1 \cdots \gamma_\omega})$, *and* $C = A \times B$. *The rank* $r(,T_1 \cdots T_p \gamma_1 \cdots \gamma_\omega; \mu_1 \cdots \mu_p \mu_{p+1} \cdots \mu_{p+\omega}; \alpha)$ *of* C *is equal to the smaller of the ranks* $r(,T_1 \cdots T_p; \mu_1 \cdots \mu_p; \alpha)$ *of* A *and* $r(,\gamma_1 \cdots \gamma_\omega; \mu_{p+1} \cdots \mu_{p+\omega}; \alpha)$ *of* B.

We shall now prove

THEOREM 30. *Let* $A = (a_{\tau_1 \cdots \tau_p})$, $B = (b_{\gamma_1 \cdots \gamma_\omega})$, $C = A \times B$. *Let* $\delta_1 = T_{\epsilon+1}\, \gamma_{\lambda+1}, \cdots, \delta_{\nu+\chi} = T_{\epsilon+\nu+\chi}\, \gamma_{\lambda+\nu+\chi}$. *Further let* $\delta = (\delta_{T_{\epsilon+1} T'_{\epsilon+1}}) \times \cdots \times (\delta_{T_{\epsilon+\nu+\chi} T'_{\epsilon+\nu+\chi}})$ *where* $(\delta_{T_{\epsilon+s} T'_{\epsilon+s}})$ *is an identity matrix on* $T_{\epsilon+s}, T'_{\epsilon+s}$ *for* $s = 1, \cdots, \nu + \chi$, *and* $\rho = T_{\epsilon+1} \cdots T_{\epsilon+\nu+\chi}$, *and let* $\delta'_1 = T'_{\epsilon+1}\, \gamma_{\lambda+1}, \cdots, \delta'_{\nu+\chi} = T'_{\epsilon+\nu+\chi}\, \gamma_{\lambda+\nu+\chi}$. *Assume that* $\epsilon \geq 1$, $\nu + \lambda \geq 1$, $\nu + \chi > 0$. *Then*

1. *if* ϵ, $\nu + \lambda$ *are both odd, the rank* $r(\delta_1 \cdots \delta_\nu T_1 \cdots T_\epsilon \gamma_1 \cdots \gamma_\lambda, T_{\epsilon+\nu+\chi+1} \cdots T_p\, \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta_{\nu+1} \cdots \delta_{\nu+\chi})$ *of* C *is less than or equal to the ranks* $r(T_1 \cdots T_\epsilon \rho, T_{\epsilon+\nu+\chi+1} \cdots T_p)$ *of* A *and* $r(\delta'_1 \cdots \delta'_\nu\, \gamma_1 \cdots \gamma_\lambda \rho, \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$ *of* $B \times \delta$.

2. *if* ϵ, $\nu + \lambda$ *are both even, the rank* $r(\delta_1 \cdots \delta_\nu T_1 \cdots T_\epsilon \gamma_1 \cdots \gamma_\lambda, T_{\epsilon+\nu+\chi+1} \cdots T_p \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta_{\nu+1} \cdots \delta_{\nu+\chi})$ *of* C *is less than or equal to the ranks* $r(T_1 \cdots T_\epsilon, \rho T_{\epsilon+\nu+\chi+1} \cdots T_p)$ *of* A *and* $r(\delta'_1 \cdots \delta'_\nu\, \gamma_1 \cdots \gamma_\lambda, \rho\gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$ *of* $B \times \delta$.

The rank $r(\delta_1 \cdots \delta_\nu T_1 \cdots T_\epsilon \gamma_1 \cdots \gamma_\lambda, T_{\epsilon+\nu+\chi+1} \cdots T_p\, \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta_{\nu+1} \cdots \delta_{\nu+\chi})$ of C is equal to the rank $r(\delta'_1 \cdots \delta'_\nu\, T_1 \cdots T_\epsilon\, \gamma_1 \cdots \gamma_\lambda, T_{\epsilon+\nu+\chi+1} \cdots T_p\, \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega\, \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$ of C_1, where

$$C_1 = A \mid \rho \mid B \times \delta.$$

According as ϵ, $\nu + \lambda$ are both odd or both even we apply Theorems 22_1 and 22_2.

For the cases where $\nu + \lambda = 0$ in the rank of C given above in Theorem 30 we prove

THEOREM 31. *Let* $A = (a_{\tau_1 \cdots \tau_p})$, $B = (b_{\gamma_1 \cdots \gamma_\omega})$, $C = A \times B$. *Let* $\chi \geq 1$. *Let* $\delta_1 = T_{\epsilon+1}\, \gamma_1, \cdots, \delta_\chi = T_{\epsilon+\chi}\, \gamma_\chi$, *and* $\delta = (\delta_{T_{\epsilon+1} T'_{\epsilon+1}}) \times \cdots \times (\delta_{T_{\epsilon+\chi} T'_{\epsilon+\chi}})$, *where* $(\delta_{T_{\epsilon+s} T'_{\epsilon+s}})$ *is an identity matrix on* $T_{\epsilon+s}, T'_{\epsilon+s}$ for $s = 1, \cdots, \chi$. *Let* $\delta'_1 = T'_{\epsilon+1}\, \gamma_1, \cdots, \delta'_\chi = T'_{\epsilon+\chi}\, \gamma_\chi$. *Further let* $\rho = T_{\epsilon+1} \cdots T_{\epsilon+\chi}$, *and let* Π *be the smallest of the orders of* $T_1, \cdots T_\epsilon$. *Then the rank* $r(T_1 \cdots T_\epsilon, T_{\epsilon+\chi+1} \cdots T_p\, \gamma_{\chi+1} \cdots \gamma_\omega\, \delta_1 \cdots \delta_\chi)$ *of* C *is less than or equal to the ranks* $r(T_1 \cdots T_\epsilon, T_{\epsilon+\chi+1} \cdots T_p \rho)$ *of* A *and* $r(,\rho\gamma_{\chi+1} \cdots \gamma_\omega\, \delta'_1 \cdots \delta'_\chi; \Pi \cdots \Pi; \Pi)$ *of* $B \times \delta$.

The rank $r(T_1 \cdots T_\epsilon, T_{\epsilon+\chi+1} \cdots T_p\, \gamma_{\chi+1} \cdots \gamma_\omega\, \delta_1 \cdots \delta_\chi)$ of C is equal to the rank $r(T_1 \cdots T_\epsilon, T_{\epsilon+\chi+1} \cdots T_p\, \gamma_{\chi+1} \cdots \gamma_\omega\, \delta'_1 \cdots \delta'_\chi)$ of C_1 where

$$C_1 = A \mid \rho \mid B \times \delta.$$

We now apply Theorem 24.

For the cases where $\epsilon = 0$ in the rank of C given in Theorem 30 we prove

THEOREM 32. *Let* $A = (a_{\tau_1 \cdots \tau_p})$, $B = (b_{\gamma_1 \cdots \gamma_\omega})$, $C = A \times B$. *Assume that* $\nu + \lambda \geq 2$, $\nu + \chi > 0$. *Let* $\delta_1 = T_1 \gamma_{\lambda+1}, \cdots, \delta_{\nu+\chi} = T_{\nu+\chi} \gamma_{\lambda+\nu+\chi}$, *and* $\delta = (\delta_{T_1 T'_1}) \times \cdots \times (\delta_{T_{\epsilon+\chi} T'_{\epsilon+\chi}})$, *where* $(\delta_{T_s T'_s})$ *is an identity matrix on* T_s, $T'_s \cdots$ *for* $s = 1, \cdots, \nu + \chi$. *Further let* $\rho = T_1 \cdots T_{\nu+\chi}$, *and let* Π *be the smallest of the orders of* $\gamma_1, \cdots, \gamma_\lambda, \delta_1, \cdots, \delta_\nu$. *Then the rank* $r(\delta_1 \cdots \delta_\nu \gamma_1 \cdots \gamma_\lambda, T_{\chi+\nu+1} \cdots T_p \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega \delta_{\nu+1} \cdots \delta_{\nu+\chi})$ *of* C *is less than or equal to the ranks* $r(\delta'_1 \cdots \delta'_\nu \gamma_1 \cdots \gamma_\lambda \rho, \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$ *of* $B \times \delta$ *and* $r(,\rho T_{\chi+\nu+1} \cdots T_p; \Pi \cdots \Pi; \Pi)$ *of* A.

The rank $r(\delta_1 \cdots \delta_\nu \gamma_1 \cdots \gamma_\lambda, T_{\chi+\nu+1} \cdots T_p \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega \delta_{\nu+1} \cdots \delta_{\nu+\chi})$ of C is equal to the rank $r(\delta'_1 \cdots \delta'_\nu \gamma_1 \cdots \gamma_\lambda, T_{\chi+\nu+1} \cdots T_p \gamma_{\lambda+\nu+\chi+1} \cdots \gamma_\omega \delta'_{\nu+1} \cdots \delta'_{\nu+\chi})$ of C_1, where

$$C_1 = A \mid \rho \mid B \times \delta.$$

We apply the Corollary to Theorem 24.

For non-signant ranks we have

THEOREM 33. *Let* $A = (a_{\tau_1 \cdots \tau_p})$, $B = (b_{\gamma_1 \cdots \gamma_\omega})$, $C = A \times B$. *Assume that* $\chi \geq 1$. *Let* $\delta_1 = T_1 \gamma_1, \cdots, \delta_\chi = T_\chi \gamma_\chi$, *and* $\delta = (\delta_{T_1 T'_1}) \times \cdots \times (\delta_{T_\chi T'_\chi})$ *where* $(\delta_{T_s T'_s})$ *is an identity matrix on* T_s, T'_s *for* $s = 1, \cdots, \chi$. *Further let* $\rho = T_1 \cdots T_\chi$. *Then the rank* $r(,T_{\chi+1} \cdots T_p \gamma_{\chi+1} \cdots \gamma_\omega \delta_1 \cdots \delta_\chi; \mu_1 \cdots \mu_{p-\chi} \beta_1 \cdots \beta_{\omega-\chi} \xi_1 \cdots \xi_\chi; \alpha)$ *of* C *is less than or equal to the ranks* $r(,\rho T_{\chi+1} \cdots T_p; \alpha\mu_1 \cdots \mu_{p-\chi}; \alpha)$ *of* A *and* $r(,\rho\gamma_{\chi+1} \cdots \gamma_\omega \delta'_1 \cdots \delta'_\chi; \alpha\beta_1 \cdots \beta_{\omega-\chi} \xi_1 \cdots \xi_\chi; \alpha)$ *of* $B \times \delta$.

The rank $r(,T_{\chi+1} \cdots T_p \gamma_{\chi+1} \cdots \gamma_\omega \delta_1 \cdots \delta_\chi; \mu_1 \cdots \mu_{p-\chi} \beta_1 \cdots \beta_{\omega-\chi} \xi_1 \cdots \xi_\chi; \alpha)$ of C is equal to the rank $r(,T_{\chi+1} \cdots T_p \gamma_{\chi+1} \cdots \gamma_\omega \delta'_1 \cdots \delta'_\chi; \mu_1 \cdots \mu_{p-\chi} \beta_1 \cdots \beta_{\omega-\chi} \xi_1 \cdots \xi_\chi; \alpha)$ of C_1, where

$$C_1 = A \mid \rho \mid B \times \delta.$$

We now apply Theorem 25.

In Theorems 27 to 33 we have established relations between all of the possible space ranks of $A \times B$ and the space ranks of A and $B \times \delta$, where δ is an identity matrix.

18. **Space ranks of Scott products.** Rice[25] has shown that the $p + q + 1$-way determinant $\mid C \mid = \mid a_{i_1 \cdots i_p T} b_{T j_1 \cdots j_q} \mid$, where T is not summed, is equal to the product of the determinants $\mid A \mid = \mid a_{i_1 \cdots i_p T} \mid$, $\mid B \mid = \mid b_{T j_1 \cdots j_q} \mid$, where the i's and j's in $\mid A \mid$ and $\mid B \mid$ have the same signancy as in $\mid C \mid$. The T's of $\mid A \mid$ and $\mid B \mid$ are both signant or both non-signant if T of $\mid C \mid$ is non-signant, while on the other hand one of the T's of $\mid A \mid$ and $\mid B \mid$ is non-signant and the other signant if T of $\mid C \mid$ is signant. We have at once

THEOREM 34. *If* $C = (a_{T_1 \cdots T_p T} b_{T \gamma_1 \cdots \gamma_\omega})$, *where* T *is not summed, and* $A = (a_{T_1 \cdots T_p T})$, $B = (b_{T \gamma_1 \cdots \gamma_\omega})$, *then*

[25] L. H. Rice, Op. Cit., p. 65.

1. *the rank* $r(T_1 \cdots T_t\, \gamma_1 \cdots \gamma_s\, T,\ T_{t+1} \cdots T_p\ \gamma_{s+1} \cdots \gamma_\omega)$, $t \geq 1$, $s \geq 1$, *of* C *is equal to the smaller of the ranks* $r(T_1 \cdots T_t\, T, T_{t+1} \cdots T_p)$ *of* A *and* $r(\gamma_1 \cdots \gamma_s, T\ \gamma_{s+1} \cdots \gamma_\omega)$ *of* B *if* t *is odd and* s *is even; the rank* $r(T_1 \cdots T_t\, \gamma_1 \cdots \gamma_s\, T, T_{t+1} \cdots T_p\, \gamma_{s+1} \cdots \gamma_\omega)$, $t \geq 1$, $s \geq 1$, *of* C *is equal to the smaller of the ranks* $r(T_1 \cdots T_t, TT_{t+1} \cdots T_p)$ *of* A *and* $r(\gamma_1 \cdots \gamma_s\, T, \gamma_{s+1} \cdots \gamma_\omega)$ *of* B *is* s *is odd and* t *is even*;

2. *the rank* $r(T_1 \cdots T_t\ \gamma_1 \cdots \gamma_s, TT_{t+1} \cdots T_p\ \gamma_{s+1} \cdots \gamma_\omega)$, $t \geq 1$, $s \geq 1$, *of* C *is equal to the smaller of the ranks* $r(T_1 \cdots T_t\, T, T_{t+1} \cdots T_p)$ *of* A *and* $r(\gamma_1 \cdots \gamma_s\, T, \gamma_{s+1} \cdots \gamma_\omega)$ *of* B *if* t, s *are odd; the rank* $r(T_1 \cdots T_t\, \gamma_1 \cdots \gamma_s, TT_{t+1} \cdots T_p\, \gamma_{s+1} \cdots \gamma_\omega)$, $t \geq 1$, $s \geq 1$, *of* C *is equal to the smaller of the ranks* $r(T_1 \cdots T_t, TT_{t+1} \cdots T_p)$ *of* A *and* $r(\gamma_1 \cdots \gamma_s, T\gamma_{s+1} \cdots \gamma_p)$ *of* B *if* t, s *are even.*

For the cases where $T_1, \cdots, T_p$ are all non-signant we have

THEOREM 35. *Let* $C = (a_{T_1 \cdots T_p T}\, b_{T\gamma_1 \cdots \gamma_\omega})$, T *not summed,* $A = (a_{T_1 \cdots T_p T})$, $B = (b_{T\gamma_1 \cdots \gamma_\omega})$. *Let* $s \geq 1$. *Let* Π *be the smallest of the orders of* $\gamma_1, \cdots, \gamma_s$ *and* Π_1 *the smallest of the orders of* $\gamma_1, \cdots, \gamma_s, T$. *Then*

1. *the rank* $r(\gamma_1 \cdots \gamma_s\, T, T_1 \cdots T_p\, \gamma_{s+1} \cdots \gamma_\omega)$ *of* C *is equal to the smaller of the ranks* $r(,TT_1 \cdots T_p; 1\Pi_1 \cdots \Pi_1; \Pi_1)$ *of* A *and* $r(\gamma_1 \cdots \gamma_s\, T, \gamma_{s+1} \cdots \gamma_\omega)$ *of* B;

2. *the rank* $r(\gamma_1 \cdots \gamma_s, TT_1 \cdots T_p\, \gamma_{s+1} \cdots \gamma_\omega)$ *of* C *is equal to the smaller of the ranks* $r(,TT_1 \cdots T_p; \Pi \cdots \Pi; \Pi)$ *of* A *and* $r(\gamma_1 \cdots \gamma_s; T\gamma_{s+1} \cdots \gamma_\omega)$ *of* B.

For the cases where no partitions are signant we have

THEOREM 36. *Let* $C = (a_{T_1 \cdots T_p T}\, b_{T\gamma_1 \cdots \gamma_\omega})$, T *not summed,* $A = (a_{T_1 \cdots T_p T})$, $B = (b_{T\gamma_1 \cdots \gamma_\omega})$. *The rank* $r(,T\gamma_1 \cdots \gamma_\omega\, T_1 \cdots T_p; \Pi\, \Pi_1 \cdots \Pi_\omega\, \Pi_{\omega+1} \cdots \Pi_{\omega+p}; \alpha)$ *of* C *is equal to the smaller of the ranks* $r(,TT_1 \cdots T_p; \Pi\, \Pi_{\omega+1} \cdots \Pi_{\omega+p}; \alpha)$ *of* A *and* $r(,T\gamma_1 \cdots \gamma_\omega; \Pi\, \Pi_1 \cdots \Pi_\omega; \alpha)$ *of* B.

All ranks of C are treated in Theorems 34 to 36 except the ranks of C which contain partitions of the type $\rho_1 = TT_1$, $\rho_2 = T_1\, T\gamma_1$, $\delta_1 = T_{\epsilon+1}\, \gamma_{\lambda+1}$, $\delta_2 = T_{\epsilon+2}\, \gamma_{\lambda+2}, \cdots$. Now the $B \times \delta$ device used to treat these ranks under zero-compositions or compositions on T_σ can not be applied to Scott products.

19. **An application of ranks of Scott Products and T_σ-composites.** We shall call the product $\left(\sum_T a_{TT_1 \cdots T_p}\, b_{T\gamma_1 \cdots \gamma_\omega}\, c_{T\sigma_1 \cdots \sigma_\rho}\right)$ of three matrices $A = (a_{TT_1 \cdots T_\rho})$, $B = (b_{T\gamma_1 \cdots \gamma_\omega})$, $C = (c_{T\sigma_1 \cdots \sigma_\rho})$, the *triple-composite* of these matrices *on* T. In another paper[26] the author has derived the conditions under which certain matrices are factorable into triple-composites. Concerning the ranks of these composites we shall prove

THEOREM 37. *Let matrices* A, B, C, D *be given by* $(a_{TT_1 \cdots T_\rho})$, $(b_{T\gamma_1 \cdots \gamma_\omega})$, $(c_{T\sigma_1 \cdots \sigma_\rho})$, $(a_{TT_1 \cdots T_p}\, b_{T\gamma_1 \cdots \gamma_\omega}\, c_{T\sigma_1 \cdots \sigma_\rho})$ *respectively. Then*

1. *if* $\epsilon, \lambda, \tau > 0$, ϵ, λ *are odd, and* τ *is even, the rank* $r(T_1 \cdots T_\epsilon \gamma_1 \cdots \gamma_\lambda \sigma_1 \cdots \sigma_\tau, T_{\epsilon+1} \cdots T_p\, \gamma_{\lambda+1} \cdots \gamma_\omega \sigma_{\gamma+1} \cdots \sigma_\rho)$ *of* D *is less than or equal to the ranks* $r(T_1 \cdots T_\epsilon T, T_{\epsilon+1} \cdots T_p)$ *of* A, $r(\gamma_1 \cdots \gamma_\lambda\, T, \gamma_{\lambda+1} \cdots \gamma_\omega)$ *of* B, *and* $r(\sigma_1 \cdots \sigma_\tau, T\sigma_{\tau+1} \cdots \sigma_\rho)$ *of* C;

[26] Equivalence of pairs of Trilinear Forms; Non-singular case, not yet published.

2. *if* $\epsilon, \lambda, \tau > 0$, *and* ϵ, λ, τ *are all even, the rank* $r(T_1 \cdots T_\epsilon \gamma_1 \cdots \gamma_\lambda \sigma_1 \cdots \sigma_\tau, T_{\epsilon+1} \cdots T_\rho \gamma_{\lambda+1} \cdots \gamma_\omega \sigma_{\tau+1} \cdots \sigma_\rho)$ *of* D *is less than or equal to the ranks* $r(T_1 \cdots T_\epsilon, TT_{\epsilon+1} \cdots T_\rho)$ *of* A, $r(\gamma_1 \cdots \gamma_\lambda, T\gamma_{\lambda+1} \cdots \gamma_\omega)$ *of* B, *and* $r(\sigma_1 \cdots \sigma_\tau, T\sigma_{\tau+1} \cdots \sigma_\rho)$ *of* C;

3. *if* $\epsilon, \tau > 0$, *and* ϵ, τ *are both odd, the rank* $r(T_1 \cdots T_\epsilon \sigma_1 \cdots \sigma_\tau, T_{\epsilon+1} \cdots T_\rho \gamma_1 \cdots \gamma_\omega \sigma_{\tau+1} \cdots \sigma_\rho)$ *of* D *is less than or equal to the ranks* $r(T_1 \cdots T_\epsilon T, T_{\epsilon+1} \cdots T_\rho)$ *of* A, $r(,T\gamma_1 \cdots \gamma_\omega; \Pi \cdots \Pi; \Pi)$ *of* B, *and* $r(\sigma_1 \cdots \sigma_\tau T, \sigma_{\tau+1} \cdots \sigma_\rho)$ *of* C, *where* Π *is the smallest of the orders of* $T_1, \cdots, T_\epsilon, \sigma_1, \cdots, \sigma_\tau$;

4. *if* $\epsilon, \tau > 0$, *and* ϵ, τ *are both even, the rank* $r(T_1 \cdots T_\epsilon \sigma_1 \cdots \sigma_\tau, T_{\epsilon+1} \cdots T_\rho \gamma_1 \cdots \gamma_\omega \sigma_{\tau+1} \cdots \sigma_\rho)$ *of* D *is less than or equal to the ranks* $r(T_1 \cdots T_\epsilon, TT_{\epsilon+1} \cdots T_\rho)$ *of* A, $r(,T\gamma_1 \cdots \gamma_\omega; \Pi \cdots \Pi; \Pi)$ *of* B, *and* $r(\sigma_1 \cdots \sigma_\tau, T\sigma_{\tau+1} \cdots \sigma_\rho)$ *of* C, *where* Π *is the smallest of the orders of* $T_1, \cdots, T_\epsilon, \sigma_1, \cdots, \sigma_\tau$;

5. *if* $\tau > 0$, *and* τ *is even, the rank* $r(\sigma_1 \cdots \sigma_\tau, T_1 \cdots T_\rho \gamma_1 \cdots \gamma_\omega \sigma_{\tau+1} \cdots \sigma_\rho)$ *of* D *is less than or equal to the ranks* $r(,TT_1 \cdots T_\rho; \Pi \cdots \Pi; \Pi)$ *of* A, $r(,T\gamma_1 \cdots \gamma_\omega; \Pi \cdots \Pi; \Pi)$ *of* B, *and* $r(\sigma_1 \cdots \sigma_\tau, T\sigma_{\tau+1} \cdots \sigma_\rho)$ *of* C, *where* Π *is the smallest of the orders of* $\sigma_1, \cdots, \sigma_\tau$;

6. *the rank* $r(,T_1 \cdots T_\rho \gamma_1 \cdots \gamma_\omega \sigma_1 \cdots \sigma_\rho; \mu_1 \cdots \mu_\rho \xi_1 \cdots \xi_\omega \beta_1 \cdots \beta_\rho; \alpha)$ *of* D *is less than or equal to the ranks* $r(,TT_1 \cdots T_\rho; \alpha\mu_1 \cdots \mu_\rho; \alpha)$ *of* A, $r(,T\gamma_1 \cdots \gamma_\omega; \alpha\xi_1 \cdots \xi_\omega; \alpha)$ *of* B, *and* $r(,T\sigma_1 \cdots \sigma_\rho; \alpha\beta_1 \cdots \beta_\rho; \alpha)$ *of* C.

Now $D = E \mid T \mid C$, where E is the Scott product $(a_{\tau\tau_1 \cdots \tau_\rho} b_{\tau\gamma_1 \cdots \gamma_\omega})$ of A and B. To prove Cases 1 to 4 of Theorem 37 we obtain inequalities between the ranks of D, E, and C by means of Theorem 22. Furthermore, for Cases 1 and 2 we obtain inequalities between the ranks of E, A, and B by using Theorem 34, and for Cases 3 and 4 by applying Theorem 35.

To treat Case 5 we obtain inequalities between the ranks of D, E, and C by using Theorem 24, and between the ranks of E, A, and B by using Theorem 36.

Finally, for Case 6 we obtain inequalities between the ranks of D, E, and C by means of Theorem 25, and inequalities between the ranks of E, A, and B by Theorem 36.

Theorem 37 covers all of the ranks of D which can be treated by means of the rank theory developed in §§15–17.

20. **Relations between the ranks of a matrix.** We shall now prove

THEOREM 38. *The rank* $r(T_1 \cdots T_\chi, T_{\chi+1} \cdots T_\xi)$, $\chi \geq 2$, *of a matrix* A *is less than or equal to the rank* $r(\rho_1 \cdots \rho_\alpha, \rho_{\alpha+1} \cdots \rho_{\alpha+\beta})$, $\alpha \geq 2$, *of* A, *where the* T*'s simply cover the* ρ*'s, and* $\rho_1, \cdots, \rho_\alpha$ *each contain an odd number of signant* T*'s, and where* $\rho_{\alpha+1}, \cdots, \rho_{\alpha+\beta}$ *each contain an even number* (≥ 0) *of signant* T*'s.*[27]

Let the rank of A in the T's be denoted by $r(T_1 \cdots T_\nu T_{\nu+1} \cdots T_{\nu+\epsilon} \cdots T_\psi \cdots T_\chi, T_{\chi+1} \cdots T_{\chi+\lambda} T_{\chi+\lambda+1} \cdots T_{\chi+\lambda+\omega} \cdots T_\mu \cdots T_\xi)$. Let $\mid M \mid_n$ denote an nth-order minor of the $(T_1 \cdots T_\chi, T_{\chi+1} \cdots T_\xi)$-derivate $(a'_{\tau_1 \cdots \tau_\chi \tau'_{\chi+1} \cdots \tau'_\xi})$ of A, and let the $(\rho_1 \cdots \rho_\alpha, \rho_{\alpha+1} \cdots \rho_{\alpha+\beta})$-derivate of A be given by

[27] The special case for Hitchcock's ranks is treated by F. L. Hitchcock, Op. Cit., p. 73. He considers relations between the ranks $r(T_1 \cdots T_{p-1}, T_p)$, $r(T_1 \cdots T_p,)$, and $r(T_1 \cdots T_r, \alpha)$, $r(T_1 \cdots T_r \alpha)$, where $\alpha = T_{r+1} \cdots T_p$, $r < p$.

$(a'_{\rho_1 \cdots \rho_\alpha \rho'_{\alpha+1} \cdots \rho'_{\alpha+\beta}})$. Let $\rho_1 = T_1 \cdots T_\nu T_{\chi+1} \cdots T_{\chi+\lambda}$, $\rho_2 = T_{\nu+1} \cdots T_{\nu+\epsilon} T_{\chi+\lambda+1} \cdots T_{\chi+\lambda+\omega}$, $\cdots$, $\rho_\alpha = T_\psi \cdots T_\chi$, $\cdots$, $\rho_{\alpha+\beta} = T_\mu \cdots T_\xi$. Let $\rho_{i1}, \rho_{i2}, \cdots, \rho_{in}$ be n values of ρ_i, $i = 1, \cdots, \alpha$, and $\rho'_{i1}, \rho'_{i2}, \cdots, \rho'_{in}$ be n values of ρ'_i for $i = \alpha + 1, \cdots, \alpha + \beta$; similarly let $T_{i1}, \cdots, T_{in}$ represent n values of T_i for $i = 1, \cdots, \chi$, and $T'_{i1}, \cdots, T'_{in}$ n values of T'_i for $i = \chi + 1, \cdots, \xi$. Let S represent the sum of all terms in the expansion of $| M |_n$ in which $T_{11}, \cdots, T_{1n}, T_{21}, \cdots, T_{2n}, \cdots, T_{\chi 1}, \cdots, T_{\chi n}, T'_{\chi+1,1}, \cdots, T'_{\chi+1,n}, \cdots, T'_{\xi 1}, \cdots, T'_{\xi n}$ occur only in the combination $\rho_{11} = T_{11} \cdots T_{\nu 1} T'_{\chi+1,1} \cdots T'_{\chi+\lambda,1}$, $\rho_{12} = T_{12} \cdots T_{\nu,2} T'_{\chi+1,2} \cdots T'_{\chi+\lambda,2}$, $\cdots$, $\rho_{1n} = T_{1n} \cdots T_{\nu n} T'_{\chi+1,n} \cdots T'_{\chi+\lambda,n}$, $\cdots$, $\rho_{\alpha+\beta,1} = T'_{\mu 1} \cdots T'_{\xi 1}$, $\cdots$, $\rho_{\alpha+\beta,n} = T'_{\mu n} \cdots T'_{\xi n}$. Let $\sigma_1, \cdots, \sigma_\alpha$ be defined by the relations

$$
\begin{aligned}
[\rho_{11} \cdots \rho_{1n}]\, \sigma_1 &= [T_{11} \cdots T_{1n}] \cdots [T_{\nu 1} \cdots T_{\nu n}], \\
[\rho_{21} \cdots \rho_{2n}]\, \sigma_2 &= [T_{\nu+1,1} \cdots T_{\nu+1,n}] \cdots [T_{\nu+\epsilon,1} \cdots T_{\nu+\epsilon,n}], \\
&\vdots \\
[\rho_{\alpha 1} \cdots \rho_{\alpha n}]\, \sigma_\alpha &= [T_{\psi 1} \cdots T_{\psi n}] \cdots [T_{\chi 1} \cdots T_{\chi n}].
\end{aligned} \tag{11}
$$

Evidently $\sigma_1, \cdots, \sigma_\alpha$ are uniquely defined by the choice of $\rho_{11}, \cdots, \rho_{1n}, \cdots, \rho_{\alpha 1}, \cdots, \rho_{\alpha n}$. Now $S = \sigma_1 \cdots \sigma_\alpha \mid M' \mid_n$ where $\mid M' \mid_n$ is the minor of the $(\rho_1 \cdots \rho_\alpha, \rho_{\alpha+1} \cdots \rho_{\alpha+\beta})$-derivate of A obtained by letting ρ_t vary over the range $\rho_{t1}, \cdots, \rho_{tn}$ for $t = 1, \cdots, \alpha$, and ρ'_t vary over $\rho'_{t1}, \cdots, \rho'_{tn}$ for $t = \alpha + 1, \cdots, \alpha + \beta$. Furthermore

$$
| M |_n = \Sigma\, \sigma_1 \cdots \sigma_\alpha \mid M' \mid_n, \tag{12}
$$

where[28] the summation is over all ρ_{ij}, $i = 1, \cdots, \alpha$; $j = 1, \cdots, n$, and ρ'_{ij}, $i = \alpha + 1, \cdots, \alpha + \beta$; $j = 1, \cdots, n$ obtained by permuting $T_{11}, \cdots, T'_{\xi n}$ in the ρ_{ij} and ρ'_{ij}. This proves the theorem.

By (11) and (12) we also have

THEOREM 39. *The rank $r(T_1 \cdots T_\chi, T_{\chi+1} \cdots T_\xi)$, $\chi \geq 2$, of a matrix A is less than or equal to the rank $r(,\rho_1 \cdots \rho_\beta; \Pi \cdots \Pi; \Pi)$ of A, where the T's simply cover the ρ's, and each of the ρ's contain an even number of signant T's, and further Π is the smallest of the orders of $T_1, \cdots T_\chi$.*

For the non-signant ranks of A we have

THEOREM 40. *Let γ_1 be the smallest of a set of positive integers $\mu_1, \cdots, \mu_\nu, \alpha; \gamma_2$ be the smallest of a set of positive integers $\mu_{\nu+1}, \cdots, \mu_{\nu+\lambda}, \alpha$; etc. Let ρ_1 be a partition of the indices of $A = (a_{T_1 \cdots T_\xi})$ which is simply covered by $T_1, \cdots, T_\nu; \rho_2$ a partition which is simply covered by $T_{\nu+1}, \cdots, T_{\nu+\lambda}$; etc. Then the rank $r(,T_1 \cdots T_\xi; \mu_1 \cdots \mu_\xi; \alpha)$ of A is less than or equal to the rank $r(,\rho_1 \cdots \rho_\beta; \gamma_1 \cdots \gamma_\beta; \alpha)$ of A.*

Let $| M |_n$, $n \leq \alpha$, denote an nth-order minor of the $(,T_1 \cdots T_\xi; \mu_1 \cdots \mu_\xi)$-

[28] The expansion (12) of $| M |_n$ is given by Rice, Op. Cit., pp. 51-54.

derivate of A. Let this derivate of A be denoted by $(a_{T_1' \cdots T_\xi'})$. Let $\rho_1' = T_1' \cdots T_\nu'$, $\rho_2' = T_{\nu+1}' \cdots T_{\nu+\lambda}'$, $\cdots$, where $\rho_1', \cdots, \rho_\beta'$ simply cover $T_1', \cdots, T_\xi'$. Let $\rho_{i1}', \cdots, \rho_{in}'$ represent n values of ρ_i' for $i = 1, \cdots, \beta$; similarly let $T_{i1}', \cdots, T_{in}'$ represent n values of T_i' for $i = 1, \cdots, \xi$. Let S represent the sum of all terms in the expansion of $| M |_n$ in which $T_{11}', \cdots, T_{1n}', T_{21}', \cdots, T_{\xi n}'$ occur only in the combination $\rho_{11}' = T_{11}' \cdots T_{\nu 1}'$, $\rho_{12}' = T_{12}' \cdots T_{\nu 2}', \cdots$. Then $S = | M' |_n$, where $| M' |_n$ is the minor of the $(,\rho_1 \cdots \rho_\beta; \gamma_1 \cdots \gamma_\beta)$-derivate $(a_{\rho_1' \cdots \rho_\beta'})$ of A obtained by letting ρ_i', $i = 1, \cdots, \beta$, vary over the range $\rho_{i1}', \cdots, \rho_{in}'$. Further $| M |_n = \Sigma | M' |_n$, where the summation extends over all ρ_{ij}', $i = 1, \cdots, \beta$; $j = 1, \cdots, n$, obtained by permuting $T_{11}', \cdots, T_{\xi n}'$ in the ρ_{ij}'. This completes the proof of the theorem.

21. **Invariance of space ranks under non-singular compositions.** If the matrix B in Theorems 22, 24, 25 is non-singular on T_σ, $\gamma = \gamma_1 \cdots \gamma_\omega$, we have the following theorem.

THEOREM 41. *Let A, B, C be given by $(a_{T_1 \cdots T_\sigma})$, $(b_{T_\sigma \gamma})$, $A \mid T_\sigma \mid B$ respectively. Let B be non-singular on T_σ, γ. Let further δ_1, ρ be defined by the relations $\delta_1 = T_{\sigma-1}\gamma$, $\rho = T_{\sigma-1}T_\sigma$. Then*

1. *the rank $r(T_1 \cdots T_r, T_{r+1} \cdots T_{\sigma-2}\delta_1)$ of C is equal to the rank $r(T_1 \cdots T_r, T_{r+1} \cdots T_{\sigma-2}\rho)$ of A;*
2. *the rank $r(T_1 \cdots T_r, T_{r+1} \cdots T_{\sigma-1}\gamma)$ of C is equal to the rank $r(T_1 \cdots T_r, T_{r+1} \cdots T_{\sigma-1}T_\sigma)$ of A;*
3. *the rank $r(T_1 \cdots T_r\,\delta_1, T_{r+1} \cdots T_{\sigma-2})$ of C is equal to the rank $r(T_1 \cdots T_r\,\rho, T_{r+1} \cdots T_{\sigma-2})$ of A;*
4. *the rank $r(T_1 \cdots T_r\,\gamma, T_{r+1} \cdots \sigma_{\sigma-1})$ of C is equal to the rank $r(T_1 \cdots T_r\,T_\sigma, T_{r+1} \cdots T_{\sigma-1})$ of A;*
5. *the rank $r(,T_1 \cdots T_{\sigma-1}\gamma; \mu_1 \cdots \mu_{\sigma-1}\alpha; \alpha)$ of C is equal to the rank $r(,T_1 \cdots T_{\sigma-1}T_\sigma; \mu_1 \cdots \mu_{\sigma-1}\alpha; \alpha)$ of A;*
6. *the rank $r(,T_1 \cdots T_{\sigma-2}\delta_1; \mu_1 \cdots \mu_{\sigma-2}\alpha; \alpha)$ of C is equal to the rank $r(,T_1 \cdots T_{\sigma-2}\rho; \mu_1 \cdots \mu_{\sigma-2}\alpha; \alpha)$ of A.*

Let $\delta = (\delta_{T_{\sigma-1}T'_{\sigma-1}})$ be an identity matrix on $T_{\sigma-1}$, $T'_{\sigma-1}$, and let $\delta_1' = T'_{\sigma-1}\gamma$. The theorem is proved by means of the relations $C = A \mid T_\sigma \mid B$, $A = C \mid \gamma \mid B^{-1}_{T_\sigma}\gamma$, Theorems 22, 24, 25, and the ranks of $C = A \mid \rho \mid B \times \delta$, $A = C \mid \delta_1' \mid (B \times \delta)^{-1}_{\rho\delta_1'}$.

22. **Space ranks of $A \times B$ where B is non-singular.** If in Theorems 27–33 B is non-singular on γ_1, γ_2, we have the following theorem.

THEOREM 42. *Let $A = (a_{T_1 \cdots T_p})$, $B = (b_{\gamma_1\gamma_2})$, $C = A \times B$. Let partitions γ, δ_0, δ_1, δ_2, ρ, ρ_1 be defined by the relations $\gamma = \gamma_1\gamma_2$, $\delta_0 = T_{s+1}\gamma_1\gamma_2$, $\delta_1 = T_{s+1}\gamma_1$, $\delta_2 = T_{s+2}\gamma_2$, $\rho = T_{s+1}T_{s+2}$. Assume that B is non-singular on γ_1, γ_2. Let Π be the order of γ_1, Π_1 the smallest of the orders of δ_1, δ_2, and let Π_2 be the smallest of the orders of $T_1, \cdots, T_s$. Then the ranks of A and C which occur in corresponding position in the following table are equal.*

A	C
$r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$	$r(T_1 \cdots T_s\gamma_1\gamma_2, T_{s+1} \cdots T_p)$, $s \geq 2$, $\Pi \geqq \Pi_2$.
$r(,T_1 \cdots T_p; \Pi \cdots \Pi; \Pi)$	$r(\gamma_1\gamma_2, T_1 \cdots T_p)$
$r(T_1 \cdots T_s, T_{s+3} \cdots T_p\rho)$	$r(T_1 \cdots T_s\delta_1\delta_2, T_{s+3} \cdots T_p)$, $s \geq 2$.
$r(,\rho T_1 \cdots T_sT_{s+3} \cdots T_p; \Pi_1 \cdots \Pi_1; \Pi_1)$	$r(\delta_1\delta_2, T_1 \cdots T_sT_{s+3} \cdots T_p)$
$r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$	$r(T_1 \cdots T_s\delta_1\gamma_2, T_{s+2} \cdots T_p)$, $s \geq 2$, $\Pi \geq \Pi_2$.
$r(,T_1 \cdots T_p; \Pi \cdots \Pi;\Pi)$	$r(\delta_1\gamma_2, T_1 \cdots T_sT_{s+2} \cdots T_p)$
$r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$	$r(T_1 \cdots T_s, \gamma_1\gamma_2\, T_{s+1} \cdots T_p)$
$r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$	$r(T_1 \cdots T_s, \gamma T_{s+1} \cdots T_p)$
$r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$	$r(T_1 \cdots T_s, T_{s+2} \cdots T_p\delta_1\gamma_2)$
$r(T_1 \cdots T_s, T_{s+3} \cdots T_p\rho)$	$t(T_1 \cdots T_s, T_{s+3} \cdots T_p\delta_1\delta_2)$
$r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$	$r(T_1 \cdots T_s, \delta_0\, T_{s+2} \cdots T_p)$
$r(T_1 \cdots T_{s+1}, T_{s+2} \cdots T_p)$	$r(T_1 \cdots T_s\delta_0, T_{s+2} \cdots T_p)$
$r(T_1 \cdots T_{s+1}, T_{s+2} \cdots T_p)$	$r(T_1 \cdots T_s\delta_1, T_{s+2} \cdots T_p\gamma_2)$
$r(T_1 \cdots T_s\rho, T_{s+2} \cdots T_p)$	$r(T_1 \cdots T_s\delta_1, T_{s+3} \cdots T_p\delta_2)$
$r(,T_1 \cdots T_p;\beta_1 \cdots \beta_p;\alpha)$	$r(,\gamma_1\gamma_2T_1 \cdots T_p; \beta\alpha\beta_1 \cdots \beta_p;\alpha)$, $\beta \geq \alpha$.
$r(,T_1 \cdots T_p;\mu_1 \cdots \mu_p;\alpha)$	$r(,\gamma T_1 \cdots T_p; \alpha\mu_1 \cdots \mu_p; \alpha)$
$r(,T_1 \cdots T_sT_{s+1}T_{s+2} \cdots T_p; \mu_1 \cdots \mu_s\alpha\mu_{s+2} \cdots \mu_p;\alpha)$	$r(,\delta_1\gamma_2T_1 \cdots T_sT_{s+2} \cdots T_p; \beta\alpha\mu_1 \cdots \mu_s\mu_{s+2} \cdots \mu_p; \alpha)$, $\beta \geq \alpha$.
$r(,\rho T_1 \cdots T_sT_{s+3} \cdots T_p; \alpha\mu_1 \cdots \mu_s\mu_{s+3} \cdots \mu_p;\alpha)$	$r(,\delta_1\delta_2T_1 \cdots T_sT_{s+3} \cdots T_p; \beta\alpha\mu_1 \cdots \mu_s\mu_{s+3} \cdots \mu_p; \alpha)$, $\beta \geq \alpha$
$r(,T_1 \cdots T_sT_{s+1}T_{s+2} \cdots T_p; \mu_1 \cdots \mu_s\alpha\mu_{s+2} \cdots \mu_p;\alpha)$	$r(,\delta_0T_1 \cdots T_sT_{s+2} \cdots T_p; \alpha\mu_1 \cdots \mu_s\mu_{s+2} \cdots \mu_p;\alpha)$

Let $\delta = (\delta_{\gamma_1\gamma_1'})$ be an identity matrix on γ_1, γ_1'; then $A \times \delta = C \mid \gamma_2 \mid B^{-1}_{\gamma_1\gamma_2}$. Let $\gamma' = \gamma_1\gamma_1'$, $\delta_0' = T_{s+1}\,\gamma_1\gamma_1'$, $\delta_2 = T_{s+2}\,\gamma_1'$. The ranks $r(\gamma_1\gamma_1', T_1 \cdots T_p)$, $r(\delta_1\delta_2', T_1 \cdots T_s, T_{s+3} \cdots T_p)$, $r(\delta_1\gamma_1', T_1 \cdots T_s\, T_{s+2} \cdots T_p)$, $r(T_1 \cdots T_s, \gamma_1\gamma_1'\, T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_s, \gamma'\, T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_s, T_{s+2} \cdots T_p\delta_1\,\gamma_1')$, $r(T_1 \cdots T_s, T_{s+3} \cdots T_p\,\delta_1\,\delta_2')$, $r(T_1 \cdots T_s, \delta_0'\, T_{s+2} \cdots T_p)$, $r(T_1 \cdots T_s\,\delta_0, T_{s+2} \cdots T_p)$, $r(T_1 \cdots T_s\delta_1, T_{s+2} \cdots T_p\gamma_1')$, $r(T_1 \cdots T_s\delta_1, T_{s+3} \cdots T_p\delta_2')$, $r(,\gamma'\, T_1 \cdots T_p; \alpha\mu_1 \cdots \mu_p; \alpha)$, $r(,\delta_0'\, T_1 \cdots T_sT_{s+2} \cdots T_p; \alpha\,\mu_1 \cdots \mu_s\mu_{s+2} \cdots \mu_p; \alpha)$ of $A \times \delta$ are respectively equal to the ranks $r(,T_1 \cdots T_p; \Pi \cdots \Pi; \Pi)$, $r(,\rho T_1 \cdots T_s\, T_{s+3} \cdots T_p; \Pi_1 \cdots \Pi_1; \Pi_1)$, $r(,T_1 \cdots T_p; \Pi \cdots \Pi; \Pi)$, $r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_s, T_{s+3}\, T_p\rho)$, $r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_{s+1}, T_{s+2} \cdots T_p)$, $r(T_1 \cdots T_{s+1}, T_{s+2} \cdots T_p)$, $r(T_1 \cdots T_s\rho, T_{s+3} \cdots T_p)$, $r(,T_1 \cdots T_p; \mu_1 \cdots \mu_p;\alpha)$, $r(,T_1 \cdots T_s\, T_{s+1}\, T_{s+2} \cdots T_p; \mu \cdots \mu_s\,\alpha\mu_{s+2} \cdots \mu_p; \alpha)$ of A. If $\beta \geq \alpha$, the ranks $r(,\gamma_1\,\gamma_1'\, T_1 \cdots T_p; \beta\,\alpha\,\beta_1 \cdots \beta_p; \alpha)$, $r(,\delta_1\gamma_1'\, T_1 \cdots T_sT_{s+2} \cdots T_p; \beta\,\alpha\mu_1 \cdots \mu_p; \alpha)$ $r(,\delta_1\delta_2'\, T_1 \cdots T_s\, T_{s+3} \cdots T_p; \beta\alpha\mu_1 \cdots \mu_s\mu_{s+3} \cdots \mu_p; \alpha)$ of $A \times \delta$ are respectively equal to the ranks $r(,T_1 \cdots T_p; \beta_1 \cdots \beta_p; \alpha)$, $r(,T_1 \cdots T_s\, T_{s+1}$

$T_{s+2} \cdots T_p;\ \mu_1 \cdots \mu_s\ \alpha\mu_{s+2} \cdots \mu_p;\ \alpha)$, $r(,\rho T_1 \cdots T_s\ T_{s+3} \cdots T_p;\ \alpha\mu_1 \cdots \mu_s\ \mu_{s+3} \cdots \mu_p;\ \alpha)$ of A. If $s \geq 2$, $\Pi \geq \Pi_2$, the ranks $r(T_1 \cdots T_s\ \gamma_1\ \gamma_1', T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_s\ \delta_1\ \gamma_1', T_{s+2} \cdots T_p)$ of $A \times \delta$ are respectively equal to the ranks $r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_s, T_{s+1} \cdots T_p)$ of A. Finally, if $s \geq 2$, the rank $r(T_1 \cdots T_s\ \delta_1\delta_2', T_{s+3} \cdots T_p)$ of $A \times \delta$ is equal to the rank $r(T_1 \cdots T_s, T_{s+3} \cdots T_p\ \rho)$ of A.

By the above relations between the ranks of A and $A \times \delta$, and Theorems 22, 24, 25, 27–33, 45, 46, Theorem 42 is proved.

From Theorem 42 several equalities may be obtained between the ranks of C.

23. **Space ranks of $C = (a_{T_1 \cdots T_p T}\ b_{T\gamma})$, T not summed, where $B = (b_{T\gamma})$ is non-singular on T, γ.** If we let $B^{-1}_{T\gamma} = (b^{T'\gamma})$, then

$$C \mid \gamma \mid B^{-1}_{T\gamma} = (a_{T_1 \cdots T_p T}\ b_{T\gamma}\ b^{T'\gamma}) = (a_{T_1 \cdots T_p T}\ \delta_{TT'}),$$

where $\delta = (\delta_{TT'})$ is an identity matrix on T, T'. We shall now prove

THEOREM 43. *Let $A = (a_{T_1 \cdots T_p T})$, $B = (b_{T\gamma})$, $C = (a_{T_1 \cdots T_p T}\ b_{T\gamma})$. If B is non-singular on T, γ, and we let Π denote the order of γ, then the ranks $r(T_1 \cdots T_s, \gamma T T_{s+1} \cdots T_p)$, $r(\gamma T, T_1 \cdots T_p)$, $r(,TT_1 \cdots T_p\gamma;\ \mu\mu_1 \cdots \mu_p\ \beta;\ \alpha)$, $\beta \geq \mu$, $r(T_1 \cdots T_s\gamma, TT_{s+1} \cdots T_p)$, $r(T_1 \cdots T_sT, \gamma T_{s+1} \cdots T_p)$ of C are equal respectively to the ranks $r(T_1 \cdots T_s, TT_{s+1} \cdots T_p)$, $r(,T_1 \cdots T_pT;\ \Pi \cdots \Pi\ 1;\ \Pi)$, $r(,TT_1 \cdots T_p;\ \mu\mu_1 \cdots \mu_p;\ \alpha)$, $r(T_1 \cdots T_sT, T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_sT, T_{s+1} \cdots T_p)$ of A.*

It is easily shown that the ranks $r(T_1 \cdots T_s, TT'T_{s+1} \cdots T_p)$, $r(T'T, T_1 \cdots T_p)$, $r(,TT_1 \cdots T_pT';\ \mu\mu_1 \cdots \mu_p\beta;\ \alpha)$, $\beta \geq \mu$, $r(T_1 \cdots T_sT', T_{s+1} \cdots T_pT)$, $r(T_1 \cdots T_sT, T'T_{s+1} \cdots T_p)$ of $(a_{T_1 \cdots T_p T}\ \delta_{TT'})$ are respectively equal to the ranks $r(T_1 \cdots T_s, TT_{s+1} \cdots T_p)$, $r(,T_1 \cdots T_pT;\ \Pi \cdots \Pi\ 1;\ \Pi)$, $r(T_1 \cdots T_pT;\ \mu_1 \cdots \mu_p\mu;\ \alpha)$, $r(T_1 \cdots T_sT, T_{s+1} \cdots T_p)$, $r(T_1 \cdots T_sT, T_{s+1} \cdots T_p)$ of A. By these relations and Theorems 35 and 36, the theorem is proved.

In general the rank $r(T_1 \cdots T_sTT', T_{s+1} \cdots T_p)$ of $(a_{T_1 \cdots T_p T}\ \delta_{TT'})$ is not greater than or equal to the rank $r(T_1 \cdots T_s, TT_{s+1} \cdots T_p)$ of A, whence the rank $r(T_1 \cdots T_sT\gamma, T_{s+1} \cdots T_p)$ of C cannot be treated by our method.

24. **Remarks on §§21–23.** It can now be shown that the restriction that the matrix B be non-singular on two partitions, where these partitions simply cover the total partition of the indices of B, can be *weakened* to non-singularity on T_σ in §21, non-singularity on γ_2 in §22, and non-singularity on T in §23. The theorems of these paragraphs can be restated in terms of these non-singularity properties.

Adopting the term *invariance* in the sense of the equality of ranks in §§21 and 22 we can state

THEOREM 44. *The signant space ranks of a matrix $A = (a_{T_1 \cdots T_p})$ are invariant under compositions on T_i with $B_i = (b_{T_i\gamma_i})$, $i = 1, \cdots, p$, where B_i is non-singular on T_i. They are also invariant under zero composition with a matrix $B = (b_{T\gamma})$, where B is non-singular on T.*

In view of the Lemma and the equivalence of a set of linear transformations to a single composition we have the

COROLLARY. *The signant ranks of a matrix A are invariant under non-singular linear transformations.*

Such invariant properties have also been established wherever possible for a large class of signant ranks under Scott multiplication, and for a large class of non-signant ranks under the different possible compositions and Scott multiplication.

25. **Ranks of $A \times B \times C \times \cdots$.** We shall prove the following generalization of Theorem 8, keeping the conventions concerning vacuous partitions adopted in §10.

THEOREM 45. *Let $\alpha = A_1 \times A_2 \times A_3 \cdots$ where $A_1, A_2, A_3, \cdots$ are matrices. Let $\rho_1, \cdots, \rho_{s+r+t+w}$ completely cover the total partition of the indices of α, and let $\rho_1 = \rho_1' \sigma_1, \cdots, \rho_s = \rho_s' \sigma_s, \rho_{s+r+1} = \rho_{s+r+1}' \sigma_{s+r+1}, \cdots, \rho_{r+s+t} = \rho_{s+r+t}' \sigma_{s+r+t}$ where $\sigma_1, \cdots, \sigma_s, \sigma_{r+s+1}, \cdots, \sigma_{r+s+t}$ are total partitions of the factor matrices $A_1, \cdots, A_s, A_{s+r+1}, \cdots, A_{s+r+t}$ respectively of α. Then the rank $r(\rho_1 \cdots \rho_s \rho_{s+1} \cdots \rho_{s+r}, \rho_{s+r+1} \cdots \rho_{s+r+t} \rho_{s+r+t+1} \cdots \rho_{s+r+t+w})$, $s + r > 0$, of α is equal to the rank $r(\rho_1' \cdots \rho_s' \rho_{s+1} \cdots \rho_{s+r}, \rho_{s+r+1}' \cdots \rho_{s+r+t}' \rho_{s+r+t+1} \cdots \rho_{s+r+t+w})$ of α', where α' is obtained from α by deleting the matrices $A_1, \cdots, A_s, A_{s+r+1}, \cdots, A_{s+r+t}$.*

If $A_1 = (a_{\sigma_1})$, then A_1 is equivalent under composition with a matrix C non-singular on σ_1 to a matrix $A_1' = (a_{\sigma_1}')$, where $a_{\sigma_1}'^* = 1$ if $\sigma_1^* = 1$, $a_{\sigma_1}'^* = 0$ if $\sigma_1^* \neq 1$. By Theorem 41 the rank $r(\rho_1 \cdots \rho_s \rho_{s+1} \cdots \rho_{s+r}, \rho_{s+r+1} \cdots \rho_{s+r+t} \rho_{s+r+t+1} \cdots \rho_{s+r+t+w})$ of α is equal to the rank $r(\rho_1' \rho_2 \cdots \rho_s \rho_{s+1} \cdots \rho_{s+r}, \rho_{s+r+1} \cdots \rho_{s+r+t} \rho_{s+r+t+1} \cdots \rho_{r+s+t+w})$ of $A_1' \times A_2 \times A_3 \cdots$, whence Theorem 45 is proved.

COROLLARY 1. *If α is not a zero matrix, and if one of the partitions $\rho_1, \cdots, \rho_s$ of Theorem 45 is the total partition of one of the factor matrices $A_1, A_2, A_3, \cdots$, then the rank $r(\rho_1 \cdots \rho_s \rho_{s+1} \cdots \rho_{s+r}, \rho_{s+r+1} \cdots \rho_{s+r+t+w})$, $s + r > 0$, of α is unity.*

COROLLARY 2. *If the partition ρ_{s+r+t} of Theorem 45 is the total partition of the factor matrix A_{s+r+t}, then the rank $r(\rho_1 \cdots \rho_{s+r}, \rho_{s+r+1} \cdots \rho_{s+r+t} \rho_{s+r+t+1} \cdots \rho_{s+r+t+w})$ of α is equal to the rank $r(\rho_1 \cdots \rho_{s+r}, \rho_{s+r+1} \cdots \rho_{s+r+t-1} \rho_{s+r+t+1} \cdots \rho_{s+r+t+w})$ of $\alpha' = A \times \cdots \times A_{s+r+t-1} \times A_{s+r+t+1} \times \cdots$.*

By a proof similar to that used for Theorem 45 we are able to prove for non-signant ranks

THEOREM 46. *Let $\alpha = A_1 \times A_2 \times A_3 \cdots$, where $A_1, A_2, A_3, \cdots$, are matrices. The rank $r(,\rho_1 \cdots \rho_s \rho_{s+1} \cdots \rho_{s+r}; \mu_1 \cdots \mu_s \mu_{s+1} \cdots \mu_{s+r}; \beta)$, $\beta = \mu_1 = \mu_2 = \cdots = \mu_s$, of α is equal to the rank $r(,\rho_1' \cdots \rho_s' \rho_{s+1} \cdots \rho_{s+r}; \mu_1 \cdots \mu_s \mu_{s+1} \cdots \mu_{s+r}; \beta)$ of α', where $\rho_1 = \rho_1' \sigma_1, \cdots, \rho_s = \rho_s' \sigma_s$, and $\sigma_1, \cdots, \sigma_s$ are total partitions of the factor matrices $A_1, \cdots, A_s$ of α; and further α' is obtained by deleting $A_1, \cdots, A_s$ from α.*

We have the

COROLLARY. *If in Theorem* 46 *the partition* ρ_1 *is the total partition of the factor matrix* A_1 *of* α, *the rank* $r(,\rho_1 \cdots \rho_s\ \rho_{s+1} \cdots \rho_{s+r};\ \mu_1 \cdots \mu_s\ \mu_{s+1} \cdots \mu_{s+r};\ \beta)$, $s + r \geq 2$, *of* α *is equal to the rank* $r(,\rho_2\rho_3 \cdots \rho_s\ \rho_{s+1} \cdots \rho_{s+r};\ \mu_2\ \mu_3 \cdots \mu_s\ \mu_{s+1} \cdots \mu_{s+r};\ \beta)$ *of* $\alpha' = A_2 \times A_3 \times \cdots$.

26. **Application of the Lemma to rank properties under compositions.** Although as mentioned in §16 the ranks of a general $B \times \delta$ are not readily expressible in terms of the ranks of B, by means of the Lemma certain 2-way ranks of $B \times \delta$ can be given in terms of the ranks of B. For the ranks of composites we have

THEOREM 47. *Let* $A = (a_{T_1 \cdots T_\omega \cdots T_{\omega+\nu} \cdots T_\sigma})$, $B = (b_{T_\sigma \gamma_1 \cdots \gamma_\omega})$, $\omega > 0$, $C = A \mid T_\sigma \mid B$. *Let* $\delta_1 = T_1\,\gamma_1, \cdots, \delta_\omega = T_\omega\,\gamma_\omega$. *Let also* $\delta = (\delta_{\gamma_1\gamma_1'}) \times (\delta_{\gamma_2\gamma_2'}) \times \cdots \times (\delta_{\gamma_{\omega-1}\gamma_{\omega-1}'})$, *where* $(\delta_{\gamma_i\gamma_i'})$, $i = 1, \cdots, \omega - 1$ *is an identity matrix on* γ_i, γ_i'. *Let* $\delta^* = (\delta_{T_\omega T_\omega'})$ *be an identity matrix on* T_ω, T_ω', *and let* $\delta_1' = T_1\,\gamma_1', \cdots, \delta_{\omega-1}' = T_{\omega-1}\,\gamma_{\omega-1}', \delta_\omega' = T_\omega'\,\gamma_\omega$. *Further let* Π *be the order of* T_ω, *and let* $\rho = T_\sigma\,\gamma_1 \cdots \gamma_{\omega-1}\,T_\omega$, $\rho_1 = T_\sigma\,\gamma_1 \cdots \gamma_{\omega-1}$. *Assume that* $\sigma > \omega$. *Then the rank* $r(\delta_1 \cdots \delta_s\ T_{\omega+1} \cdots T_{\omega+\nu}\delta_\omega,\ \delta_{s+1} \cdots \delta_{\omega-1}\ T_{\omega+\nu+1} \cdots T_{\sigma-1})$ *of* C *is less than or equal to the rank* $r(\delta_1' \cdots \delta_s'\ T_{\omega+1} \cdots T_{\omega+\nu}\ \rho,\ \delta_{s+1}' \cdots \delta_{\omega-1}'\ T_{\omega+\nu+1} \cdots T_{\sigma-1})$ *of* $A \times \delta$ *and* Πr, *where* r *is the rank* $r(\rho_1\,\gamma_\omega,)$ *of* B.

The rank $r(\delta_1 \cdots \delta_s\ T_{\omega+1} \cdots T_{\omega+\nu}\ \delta_\omega,\ \delta_{s+1} \cdots \delta_{\omega-1}\ T_{\omega+\nu+1} \cdots T_{\sigma-1})$ of C is equal to the rank $r(\delta_1' \cdots \delta_s'\ T_{\omega+1} \cdots T_{\omega+\nu}\ \delta_\omega', \delta_{s+1}' \cdots \delta_{\omega-1}'\ T_{\omega+\nu+1} \cdots T_{\sigma-1})$ of C_1, where

$$C_1 = (A \times \delta) \mid \rho \mid (B \times \delta^*).$$

The theorem now follows by means of Theorem 22 and the Lemma.

For the ranks of zero-composites we have

THEOREM 48. *Let* $A = (a_{T_1 \ldots T_\omega \ldots T_{\omega+\nu} \ldots T_\sigma})$, $B = (b_{\gamma_1 \ldots \gamma_\omega})$, $C = A \times B$. *Let* $\delta_1 = T_1\,\gamma_1, \cdots, \delta_\omega = T_\omega\,\gamma_\omega$, *and let* $\delta = (\delta_{\gamma_1\gamma_1'}) \times \cdots \times (\delta_{\gamma_{\omega-1}\gamma_{\omega-1}'})$, *where* $(\delta_{\gamma_i\gamma_i'})$, $i = 1, \cdots, \omega - 1$ *is an identity matrix on* γ_i, γ_i'. *Let* $\delta^* = (\delta_{T_\omega T_\omega'})$ *be an identity matrix on* T_ω, T_ω'. *Let also* $\delta_1' = T_1\,\gamma_1', \cdots, \delta_{\omega-1}' = T_{\omega-1}\,\gamma_{\omega-1}', \delta_\omega' = T_\omega'\,\gamma_\omega$. *Let* Π *be the order of* T_ω, *and let* $\rho = \gamma_1 \cdots \gamma_{\omega-1}\,T_\omega$, $\rho_1 = \gamma_1 \cdots \gamma_{\omega-1}$. *Then the rank* $r(\delta_1 \cdots \delta_s\ T_{\omega+1} \cdots T_{\omega+\nu}\ \delta_\omega,\ \delta_{s+1} \cdots \delta_{\omega-1}\ T_{\omega+\nu+1} \cdots T_{\sigma-1})$ *of* C *is less than or equal to the rank* $r(\delta_1' \cdots \delta_s'\ T_{\omega+1} \cdots T_{\omega+\nu}\ \rho,\ T_{\omega+\nu+1} \cdots T_{\sigma-1})$ *of* $A \times \delta$ *and* Πr, *where, if* $\omega > 1$, r *is the rank* $r(\rho_1\,\gamma_\omega,)$ *of* B, *while if* $\omega = 1$, $r = 1$ *or* 0 *according as* $B \neq 0$ *or* $B = 0$.

To prove this theorem we construct a matrix C_1 as in the proof of Theorem 47; then apply the Lemma and Theorem 22.

The ranks covered in Theorems 47 and 48 are the most general ranks which can be treated by applying the Lemma to matrices of the type $B \times \delta$.

27. **Space determinants as relative invariants.** Theorem 20 and the Lemma imply

THEOREM 49. *If a matrix* $A = (a_{T_1 \ldots T_p})$ *possesses a determinant* $\mid a_{T_1 \ldots T_p} \mid$

signant on $T_1, T_2, \cdots, T_r$, then the corresponding determinant of C obtained from A under composition with $B_{ij} = (b_{T^i_{i1} T^i_{i2}})$ on T^i_{i1}, $i = 1, 2, \cdots, n_j$, simply covering T_j, $j = 1, \cdots, r$, is equal to

$$a_{T_1 \cdots T_p} \,|\, |\, b_{T^1_{11}\, T^1_{12}} \,|^{\Pi^1_2 \cdots \Pi^1_{n_1}} \cdots |\, b_{T^1_{n_1 1}\, T^1_{n_1 2}} \,|^{\Pi^1_1 \cdots \Pi^1_{n_1 - 1}}$$

$$\cdots \,|\, b_{T^r_{n_r 1}\, T^r_{n_r 2}} \,|^{\Pi^r_1 \cdots \Pi_{n_r - 1}},$$

where Π^i_i is the order of T^i_{i1}.

The notations of the theorem imply that the matrices $(b_{T^i_{i1} T^i_{i2}})$ are square as displayed.

Relative invariant properties under compositions do not hold for determinants non-signant on all indices, or for compositions on non-signant indices of determinants signant on two or more indices.

28. **Space ranks of identity matrices.** We shall prove

THEOREM 50. *Let $\delta = (\delta_{i_1 \cdots i_r j_1 \cdots j_r}) = (\delta_{i_1 j_1}) \times \cdots \times (\delta_{i_r j_r})$, where $(\delta_{i_s j_s})$ is an identity matrix on i_s, j_s for $s = 1, \cdots, r$. Let also $\delta = (\delta_{\rho_1 \cdots \rho_w \rho'_1 \cdots \rho'_w}) = (\delta_{\rho_1 \rho'_1}) \times \cdots \times (\delta_{\rho_w \rho'_w})$, where the ρ's are simply covered by $i_1, \cdots, i_r, j_1, \cdots, j_r$, and $(\delta_{\rho_s \rho'_s})$ is an identity matrix on ρ_s, ρ'_s for $s = 1, \cdots, w$. Let $\sigma_1, \cdots, \sigma_{m+l}$ be simply covered by the j indices, and let $\rho = i_1 \cdots i_r$. Let also $i_1, \cdots, i_r$, $j_1, \cdots, j_r$ simply cover partitions $\beta_1, \cdots, \beta_p$. Then*

1. *the rank $r(\beta_1 \cdots \beta_t, \beta_{t+1} \cdots \beta_p)$, $t \geq 2$, of δ is unity if one of the signant partitions $\beta_1, \cdots, \beta_t$ contains only subsets of the type $i_a j_a$;*

2. *the rank $r(\rho_1 \rho'_1\ \rho_2 \rho'_2 \cdots \rho_j \rho_{j+1}, \cdots, \rho'_j \rho'_{j+1} \cdots \rho_w \rho'_w)$ of δ is equal to one if partitions $\rho_j, \rho_{j+1}, \cdots$, exist among the signant partitions for which $\rho'_j, \rho'_{j+1}, \cdots$, occur among the non-signant partitions; otherwise this rank is equal to the smallest of the orders of the partitions $\rho_1, \cdots, \rho_w$;*

3. *the rank $r(\rho\sigma_1 \cdots \sigma_m, \sigma_{m+1} \cdots \sigma_{m+l})$ of δ is equal to the smallest of the orders of $\sigma_1, \cdots, \sigma_{m+l}$; also the rank $r(\sigma_1 \cdots \sigma_m, \sigma_{m+1} \cdots \sigma_{m+l}\ \rho)$ of δ is equal to the smallest of the orders of the partitions $\sigma_1, \cdots, \sigma_{m+l}$;*

4. *the rank $r(,\rho_1 \rho'_1 \cdots \rho_w \rho'_w; \mu_1 \mu'_1 \cdots \mu_w \mu'_w; \alpha)$ of δ is equal to α.*

Part 1 of the theorem follows from Cor. 1 of Theorem 45.

To prove part 2 consider a qth-order ($q \geq 2$) minor of $(\delta_{\rho_1 \cdots \rho_w})$ given by the determinant $d = |\, \delta_{\rho_1 \rho'_1 \rho_2 \rho'_2 \cdots \rho_j \rho_{j+1} \cdots \rho_l, \rho'_j \rho'_{j+1} \cdots \rho_w \rho'_w} \,|$, where the partitions to the left of the comma in d are signant and ρ_i, ρ'_i vary over the same ranges. Let $\rho_{i1}, \cdots, \rho_{iq}$ denote q values of ρ_i, $i = 1, \cdots, w$; similarly for ρ'_i. The sign of $\rho_1, \rho'_1, \rho_2, \rho'_2, \cdots$, does not affect the sign of the terms in the expansion of d since $[\rho_{11} \cdots \rho_{1q}]\ [\rho'_{11} \cdots \rho'_{1q}] = [\rho_{21} \cdots \rho_{2q}]\ [\rho'_{21} \cdots \rho'_{2q}] = \cdots = 1$. Hence the expansion of d is a sum of positive terms of the type d_1, where $d_1 = |\, d_{\rho_j \rho_{j+1} \cdots} \,|$, and $d_{\rho_j \rho_{j+1}} = 1$ for all values of $\rho_j, \rho_{j+1}, \cdots$; further all partitions exhibited in d_1 are signant, whence $d_1 = 0$.

If there are no partitions of the type $\rho_j, \rho_{j+1}, \cdots$, in d, then all terms in the expansion of d are positive and $d \neq 0$.

To prove part 3 let Π be the smallest of the orders of $\sigma_1, \cdots, \sigma_{m+l}$. Let a minor of $(\delta_{\rho\sigma 1} \cdots {}_{\sigma m+l})$ be given by the determinant $| E | = | \delta_{\rho\sigma 1} \cdots {}_{\sigma m+l} |$, where σ_i varies over a range $\sigma_{i1}, \cdots, \sigma_{i\Pi}$ for $i = 1, \cdots, m + l$, and ρ varies over the range $\sigma_{11}\sigma_{21} \cdots \sigma_{m+l,\,1}, \cdots, \sigma_{1\Pi}\sigma_{2\Pi} \cdots \sigma_{m+l,\Pi}$. For every non-vanishing term in the expansion of $| E |$ the following relations hold:

$$[\rho_{11} \cdots \rho_{1q}] = [\sigma_{11} \cdots \sigma_{1q}] = \cdots = [\sigma_{m1} \cdots \sigma_{mq}]. \tag{13}$$

Since there are an even number of partitions signant (13) implies that $| E | \neq 0$ We can similarly show that the determinant $| E |$ expanded with $\rho, \sigma_{m+1}, \cdots, \sigma_{m+l}$ non-signant and $\sigma_1, \cdots, \sigma_{m-1}, \sigma_m$ signant is not zero.

Finally consider a minor of the $(,\rho_1\rho_1' \cdots \rho_w\rho_w'; \mu_1\mu_1' \cdots \mu_w\mu_w')$-derivate of δ obtained by letting ρ_j, ρ_j' have the same range for $j = 1, \cdots, w$, where the ρ's have α values. Since all terms in the expansion of this minor are positive, it is not zero. This proves part 4 of the theorem.

29. **Relation between the 3-way and 2-way ranks of a matrix.** The proof of the following theorem is a modification of that given by Hitchcock[29] for triadics.

THEOREM 51. *Let partitions $\alpha, \beta, \rho, T, \sigma$ be related by the identities $\alpha = T\sigma$, $\beta = \rho\sigma$. If the rank $r(\rho T, \sigma)$ of the matrix $(a_{\rho T\sigma})$ is unity, then one of the ranks $r(\rho\alpha,), r(T\beta,)$ of $(a_{\rho T\sigma})$ is unity.*

By assumption, if ρ_1, ρ_2 denote different values of ρ, T_1, T_2 different values of T, and σ_1, σ_2 different values of σ, then

$$\sum_{\substack{(\rho_1 \rho_2) \\ (T_1 T_2)}} [\rho_1\rho_2]\,[T_1T_2]\, a_{\rho_1\sigma_1 T_1}\, a_{\rho_2\sigma_2 T_2} = 0 \tag{14}$$

for every choice of $\rho_1, \rho_2, T_1, T_2, \sigma_1, \sigma_2$. If $\sigma_1 = \sigma_2$, (14) implies that $a_{\rho T\sigma} = \alpha^{\sigma}_{\rho}\beta^{\sigma}_{T}$. Substituting in (14) with $\sigma_1 \neq \sigma_2$, we get

$$\alpha^{\sigma_1}_{\rho_1}\beta^{\sigma_1}_{T_1}\alpha^{\sigma_2}_{\rho_2}\beta^{\sigma_2}_{T_2} - \alpha^{\sigma_1}_{\rho_2}\beta^{\sigma_1}_{T_1}\alpha^{\sigma_2}_{\rho_1}\beta^{\sigma_2}_{T_2} + \alpha^{\sigma_1}_{\rho_2}\beta^{\sigma_1}_{T_2}\alpha^{\sigma_2}_{\rho_1}\beta^{\sigma_2}_{T_1} - \alpha^{\sigma_1}_{\rho_1}\beta^{\sigma_1}_{T_2}\alpha^{\sigma_2}_{\rho_2}\beta^{\sigma_2}_{T_1} = 0, \tag{15}$$

whence we have

$$(\alpha^{\sigma_1}_{\rho_1}\alpha^{\sigma_2}_{\rho_2} - \alpha^{\sigma_1}_{\rho_2}\alpha^{\sigma_2}_{\rho_1})\,(\beta^{\sigma_1}_{T_1}\beta^{\sigma_2}_{T_2} - \beta^{\sigma_1}_{T_2}\beta^{\sigma_2}_{T_1}) = 0. \tag{16}$$

Now the second order minors of $(a_{\rho\alpha})$ are given by

$$\beta^{\sigma_1}_{T_1}\beta^{\sigma_2}_{T_2}\begin{vmatrix} \alpha^{\sigma_1}_{\rho_1} & \alpha^{\sigma_2}_{\rho_1} \\ \alpha^{\sigma_1}_{\rho_2} & \alpha^{\sigma_2}_{\rho_2} \end{vmatrix},$$

---

[29] F. L. Hitchcock, Op. Cit., pp. 70 ff.

The theorem implies the condition $| \sigma U + \tau W | \not\equiv 0$, which occurs in the author's paper "On canonical binary trilinear forms," Bulletin of the American Mathematical Society, vol. 38 (1932), p. 387.

while the corresponding minors of $(a_{T\beta})$ are given by

$$\alpha^{\sigma}_{\rho 1}\,\alpha^{\sigma}_{\rho 2}\begin{vmatrix}\beta^{\sigma}_{T1} & \beta^{\sigma}_{T2}\\ \beta^{\sigma}_{T2} & \beta^{\sigma}_{T2}\end{vmatrix}.$$

Assume that the rank $r(\rho\alpha,) > 1$, whence some minor of $(a_{\rho\alpha})$ is not zero. Let this minor be given by $\beta^{s}_{t1}\,\beta^{s}_{t2}\,(\alpha^{s}_{r1}\,\alpha^{s}_{r2} - \alpha^{s}_{r2}\,\alpha^{s}_{r1})$. Now at least one of the products $\alpha_{r1t1s1}\,a_{r2t2s2}$, $a_{r1t2s2}\,a_{r2t1s1}$ is not zero. In our proof we shall assume that $a_{r1t1s1}\,a_{r2t2s2} \neq 0$

We wish to prove that $r(T\beta,) = 1$. It is evident that it is sufficient to prove only that all second order minors of $(a_{T\beta})$ which contain a_{r1t1s1} are zero, i.e. that $\alpha^{s}_{r1}\,\alpha^{\sigma}_{\rho}\,(\beta^{s}_{t1}\,\beta^{\sigma}_{T} - \beta^{s}_{T1}\,\beta^{\sigma}_{t1}) = 0$ for all ρ, T, σ; or what is the same, that

$$\alpha^{\sigma}_{\rho}\,(\beta^{s}_{t1}\,\beta^{\sigma}_{T} - \beta^{s}_{T1}\,\beta^{\sigma}_{t1}) = 0 \tag{17}$$

for all ρ, T, σ.

If $\alpha^{s3}_{\rho} = 0$ for all ρ, or if $\beta^{s3}_{T} = 0$ for all T, (17) is satisfied. We hence assume that $a_{r3t3s3} = \alpha^{s3}_{r3}\,\beta^{s3}_{t3} \neq 0$. By (16) we have

$$(\alpha^{s1}_{r1}\,\alpha^{s2}_{r2} - \alpha^{s1}_{r2}\,\alpha^{s2}_{r1})\,(\beta^{s1}_{t1}\,\beta^{s2}_{T} - \beta^{s1}_{T1}\,\beta^{s2}_{t1}) = 0, \tag{18}$$

$$(\alpha^{s1}_{r1}\,\alpha^{s3}_{r3} - \alpha^{s1}_{r3}\,\alpha^{s3}_{r1})\,(\beta^{s1}_{t1}\,\beta^{s3}_{T} - \beta^{s1}_{T1}\,\beta^{s3}_{t1}) = 0, \tag{19}$$

$$(\alpha^{s2}_{r1}\,\alpha^{s3}_{r2} - \alpha^{s2}_{r2}\,\alpha^{s3}_{r1})\,(\beta^{s2}_{t1}\,\beta^{s3}_{T} - \beta^{s2}_{T1}\,\beta^{s3}_{t1}) = 0. \tag{20}$$

By (18) we have

$$(\beta^{s1}_{t1}\,\beta^{s2}_{T} - \beta^{s1}_{T1}\,\beta^{s2}_{t1}) = 0. \tag{21}$$

If $\beta^{s2}_{t1} \neq 0$, (20) and (21) give

$$(\alpha^{s2}_{r1}\,\alpha^{s3}_{r2} - \alpha^{s2}_{r2}\,\alpha^{s3}_{r1})\,(\beta^{s1}_{t1}\,\beta^{s3}_{T} - \beta^{s1}_{T1}\,\beta^{s3}_{t1}) = 0. \tag{22}$$

If also $\alpha^{s3}_{r2} \neq 0$, and $(\beta^{s1}_{t1}\,\beta^{s3}_{T} - \beta^{s1}_{T1}\,\beta^{s3}_{t1}) \neq 0$, then the equation

$$(\alpha^{s1}_{r1}\,\alpha^{s3}_{r2} - \alpha^{s1}_{r2}\,\alpha^{s3}_{r1}) = (\alpha^{s2}_{r1}\,\alpha^{s3}_{r2} - \alpha^{s2}_{r2}\,\alpha^{s3}_{r1}) = 0, \tag{23}$$

which follows from (19) and (22), contradicts the assumption that $(\alpha^{s1}_{r1}\,\alpha^{s2}_{r2} - \alpha^{s1}_{r2}\,\alpha^{s2}_{r1}) \neq 0$. Hence $\beta^{s1}_{t1}\,\beta^{s3}_{T} - \beta^{s1}_{T1}\,\beta^{s3}_{t1} = 0$.

If $\beta^{s2}_{t1} \neq 0$, $\alpha^{s3}_{r2} = 0$, and further $(\beta^{s1}_{t1}\,\beta^{s3}_{T} - \beta^{s1}_{T1}\,\beta^{s3}_{t1}) \neq 0$, by (22) $\alpha^{s3}_{r1} = 0$. Now from (16) we have

$$(\alpha^{s1}_{r1}\,\alpha^{s3}_{\rho} - \alpha^{s1}_{\rho}\,\alpha^{s3}_{r1})\,(\beta^{s1}_{t1}\,\beta^{s3}_{T} - \beta^{s1}_{T1}\,\beta^{s3}_{t1}) = 0 \tag{24}$$

for all ρ. This implies that $\alpha^{s3}_{\rho} = 0$ for all ρ, which contradicts the assumption that $a_{r3t3s3} \neq 0$.

If finally $\beta^{s2}_{t1} = 0$, by (21) $\beta^{s2}_{T} = 0$ for all T, which implies that $\beta^{s2}_{t2} = 0$, giving a contradiction.

30. **Compound matrices.** Let $A = (a_{T_1 \dots T_\sigma})$, where $T_1, \cdots, T_\sigma$ are each of order n. Using double subscripts on partitions in the sense of Section 14, we define a form

$$F(A) = \sum_{\substack{(T_{11} \cdots T_{1n}) \\ \vdots \\ (T_{\sigma-1,1} \cdots T_{\sigma-1,n})}} [T_{11} \cdots T_{n1}] \cdots [T_{\sigma-1,1} \cdots T_{\sigma-1,n}]\, a_{T_{11} \dots T_{\sigma 1}} \cdots a_{T_{1q} \dots T_{\sigma q}}\, z^{(1)}_{T_{\sigma 1}} \cdots z^{(q)}_{T_{\sigma q}}\, y^{(q+1)}_{T_{1,q+1}}\, v^{(q+1)}_{T_{2,q+1}} \cdots u^{(q+1)}_{T_{\sigma-1,q+1}} \cdots y^{(n)}_{T_{1n}}\, v^{(n)}_{T_{2,n}} \cdots u^{(n)}_{T_{\sigma-1,n}}.$$

The elements of the matrix A_1 of $F(A)$ are the signed determinants of the qth order minors of the $(T_1 \cdots T_{\sigma-1}, T_\sigma)$-derivate of A. A_1 is called the $(T_1 \cdots T_{\sigma-1}, T_\sigma)$-*compound*[30] of A. Let further

$$F(A \mid T_1 \mid B) = \sum_{\substack{(\rho_{11} \cdots \rho_{1n}) \\ (T_{21} \cdots T_{2n}) \\ \vdots \\ (T_{\sigma-1,1} \cdots T_{\sigma-1,n})}} [\rho_{11} \cdots \rho_{1n}]\, [T_{21} \cdots T_{2n}] \cdots [T_{\sigma-1,1} \cdots T_{\sigma-1,n}]\, a_{T_{11} \dots T_{\sigma 1}} \cdots a_{T_{1q} \dots T_{\sigma q}}\, b_{T_{11}\rho_{11}} \cdots b_{T_{1q}\rho_{1q}}\, z^{(1)}_{T_{\sigma 1}} \cdots z^{(q)}_{T_{\sigma q}}\, w^{(q+1)}_{\rho_{1,q+1}} \cdots w^{(n)}_{\rho_{1n}}\, v^{(q+1)}_{T_{2,q+1}} \cdots u^{(q+1)}_{T_{\sigma-1,q+1}}\, v^{(n)}_{T_{2n}} \cdots u^{(n)}_{T_{\sigma-1,n}}.$$

If $q < n$, $F(A \mid T_1 \mid B)$ can be obtained except for a constant factor ($\neq 0$) from $F(A)$ by the following composition

$$(25)\quad y^{(q+1)}_{T_{1,q+1}} \cdots y^{(n)}_{T_{1n}} = \sum_{\substack{(T_{11} \cdots T_{1q}) \\ (\rho_{11} \cdots \rho_{1q})}} [T_{11} \cdots T_{1n}]\, [\rho_{11} \cdots \rho_{1n}]\, b_{T_{11}\rho_{11}} \cdots b_{T_{1q}\rho_{1q}}\, w^{(q+1)}_{\rho_{1,q+1}} \cdots w^{(n)}_{\rho_{1n}}.$$

The coefficients $\sum_{\substack{(T_{11} \cdots T_{1q}) \\ (\rho_{11} \cdots \rho_{1q})}} [T_{11} \cdots T_{1n}]\, [\rho_{11} \cdots \rho_{1n}]\, b_{T_{11}\rho_{11}} \cdots b_{T_{1q}\rho_{1q}}$ are signed minors of $B = (b_{T_1\rho_1})$. The equations (25) represent a non-singular composition on $T_{1,q+1} \cdots T_{1n}$ if B is non-singular[31] on T_1, ρ_1. If $q = n$, we let[32]

$$(26_1)\qquad z^{(1)}_{T_{\sigma 1}} \to \Delta\, z^{(1)}_{T_{\sigma 1}},$$

$$(26_2)\qquad z^{(2)}_{T_{\sigma 2}} \to z^{(2)}_{T_{\sigma 2}},$$

$$\vdots \qquad\qquad \vdots$$

$$(26_q)\qquad z^{(q)}_{T_{\sigma q}} \to z^{(q)}_{T_{\sigma q}},$$

[30] For two-way compounds see Cullis, "Matrices and Determinoids," vol. 1, p. 289.

[31] Cullis, Ibid. This property is used by J. Williamson in his paper entitled "Matrices whose sth compounds are equal," Bulletin of the American Mathematical Society, vol. 39 (1933), p. 108.

[32] "$\to$" means "go into."

where $\Delta = |\, b_{T_1\rho_1} \,|$, and the determinants of the transformations (26_1), (26_2), $\cdots$, (26_q) are Δ^n, $1, \cdots, 1$ respectively. This completes the case where a composition is made on a signant partition of $F(A)$.

To treat the remaining case let $C = (c_{T_\sigma\rho_1})$. Let further

$$F(A|\, T_\sigma \,|\, C) = \sum_{\substack{(T_{11}\cdots T_{1n}) \\ \vdots \\ (T_{\sigma-1,1}\cdots T_{\sigma-1,n})}} [T_{11}\ \cdots\ T_{1n}]\ \cdots\ [T_{\sigma-1,1}\ \cdots\ T_{\sigma-1,n}]\ a_{T_{11}\ldots T_{\sigma 1}} \cdots a_{T_{1q}\ldots T_{\sigma q}}\, c_{T_{\sigma 1}\rho_{11}} \cdots c_{T_{\sigma p}\rho_{1q}}\, y^{(q+1)}_{T_{1,q+1}} \cdots u^{(q+1)}_{T_{\sigma-1,q+1}} \cdots y^{(n)}_{T_{1n}} \cdots u^{(n)}_{T_{\sigma-1,n}}\, w^{(1)}_{\rho_{11}} \cdots w^{(q)}_{\rho_{q1}}.$$

$F(A \mid T_\sigma \mid C)$ can be obtained from $F(A)$ by the compositions

$$\begin{aligned} z^{(1)}_{T_{\sigma 1}} &= \sum_{\rho_{11}} c_{T_{\sigma 1}\rho_{11}}\, w^{(1)}_{\rho_{11}}, \\ &\cdots\cdots\cdots \\ z^{(q)}_{T_{\sigma q}} &= \sum_{\rho_{1q}} c_{T_{\sigma q}\rho_{1q}}\, w^{(q)}_{\rho_{1q}}, \end{aligned} \tag{27}$$

where C is the matrix of composition in each of the equations of (27). If C is non-singular on T_σ, ρ_1, the above compositions are non-singular on $T_{\sigma 1}$, ρ_{11}, $\cdots$, $T_{\sigma q}$, ρ_{1q} respectively.

The above treatment can be extended at once to the cases where $T_1, \cdots, T_\sigma$ are not of the same orders, and the non-singularity of B and C on T_1, ρ_1 and T_σ, ρ_1 respectively can be replaced by non-singularity on T_1 and T_σ. We have now proved that compositions of A on T_i, or on partitions contained in T_i $(i = 1, \cdots, \sigma)$, with matrices non-singular on these partitions, correspond to compositions on the $(T_1 \cdots T_{\sigma-1}, \tau_\sigma)$-compound of A with matrices non-singular on the partitions on which these compositions are made; hence the space ranks of the $(T_1 \cdots T_{\sigma-1}, T_\sigma)$-compound of A are invariant, in the sense of Theorem 41, under non-singular compositions on A.

Nothing further can be obtained by using derivates with more than one partition signant.

31. **Invariant factors.** Let $A = (a_{i_1 \ldots i_n})$ be a cube matrix of order m on each index. Let $\delta = (\delta_{i_1 \ldots i_n})$, where $\delta_{i_1 \ldots i_n} = 1$ when $i_1 = \cdots = i_n = 1, \cdots, m$, and $\delta_{i_1 \ldots i_n} = 0$ for all other values of the subscripts. Let $A_1 = (a^{(1)}_{i_1 j_1})$, $\cdots$, $A_n = (a^{(n)}_{i_n j_n})$ be non-singular square matrices which satisfy

$$(\delta_{i_1 \ldots i_n}\, a^{(1)}_{i_1 j_1} \cdots a^{(n)}_{i_n j_n}) = (\delta_{j_1 \ldots j_n}) = (\delta_{i_1 \ldots i_n}). \tag{28}$$

Let simultaneous transformations on δ with $A_1, A_2, \cdots, A_n$ be called a *similar* transformation. Let λ represent a parameter. By Theorems 20 and 21 the greatest common divisor of the qth-order minors of the $(i_1 \cdots i_s, i_{s+1} \cdots i_n)$-derivate of $(A - \lambda\delta)$ of any signancy is equal to the greatest common divisor of

the minors of the corresponding derivate of $(B - \lambda\delta)$, where B is obtained from A under a similar transformation. The greatest common divisor of the determinants of the rth-order minors of the $(i_1 \cdots i_s, i_{s+1} \cdots i_n)$-derivate of $(A - \lambda\delta)$ are divisible by the greatest common divisor of the determinants of the $(r-1)$st order minors of this derivate. Let the quotient of these greatest common divisors be called the rth-*invariant factor* of the $(i_1 \cdots i_s, i_{s+1} \cdots i_n)$-derivate of A. We similarly define the rth-invariant factor of the $(,i_1 \cdots i_n; \alpha \cdots \alpha)$-derivate of A.

We have

THEOREM 52. *The invariant factors of a matrix A are unchanged under similar transformations.*

It is readily shown that except in the two-way case, the invariant factors do not form a complete invariant system under similar transformations. In another paper the author has given such a complete system of invariants.

CLEVELAND, OHIO.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

COMPOSITION AND RANK OF N-WAY MATRICES AND MULTILINEAR FORMS—SUPPLEMENT[1]

BY RUFUS OLDENBURGER

(Received January 18, 1934)

1. **Introduction.** Where the composite of two or more matrices is made by summing on at least one combined index and not summing on at least one other combined index we have the *multiple composite* of these matrices. In this supplement we obtain certain relations between the space ranks of the multiple composite of a set of matrices, obtained by combining indices in any way, and the space ranks of these matrices.

As an example of the multiple composite of a set of matrices we have the composite

$$(a_{iklmn_1}b_{ijklmn_2}c_{ijkln_3}) \qquad (k, l, m \text{ not summed; } i, j \text{ summed}),$$

of the matrices (a_{iklmn_1}), (b_{ijklmn_2}), (c_{ijkln_3}). Also the composite

$$(a_{ikm}b_{ir}c_{kpt}d_{pts}) \qquad (i, t \text{ not summed; } k, p \text{ summed}),$$

of the matrices (a_{ikm}), (b_{ir}), (c_{kpt}), (d_{pts}).

2. **Definitions and notations.** If a partition τ is given by $\tau = \rho\sigma \cdots$ we say that ρ *partly covers* and is *contained in* τ. For brevity, a set of partitions $\tau_1, \cdots, \tau_\epsilon$ will be denoted by $[\tau]$. The partitions $\tau_1, \cdots, \tau_\epsilon$ are said to be contained in $[\tau]$. A set $[\tau]$ is said to be vacuous if it contains no partitions. We shall write the rank $r(\tau_1 \cdots \tau_\rho, \tau_{\rho+1} \cdots \tau_p)$ as $r([\tau], [\tau'])$, where $[\tau] = \tau_1, \cdots, \tau_\rho$; $[\tau'] = \tau_{\rho+1}, \cdots, \tau_p$; also $r(, \tau_1 \cdots \tau_p; \pi_1 \cdots \pi_p; \alpha)$ as $r(, [\tau]; [\pi]; \alpha)$, where $[\tau] = \tau_1, \cdots, \tau_p$, and $[\pi]$ is the set of integers $\pi_1, \cdots \pi_p$. Finally we shall write $r(\tau_1 \cdots \tau_\epsilon\tau_{\epsilon+1} \cdots \tau_\rho, \tau_{\rho+1} \cdots \tau_p)$ as $r([\tau'] [\tau''], [\tau'''])$, where $[\tau'] = \tau_1, \cdots, \tau_\epsilon$, $[\tau''] = \tau_{\epsilon+1}, \cdots, \tau_\rho$, $[\tau'''] = \tau_{\rho+1}, \cdots, \tau_p$; and $r(, \tau_1 \cdots \tau_\epsilon\tau_{\epsilon+1} \cdots \tau_p; \pi_1 \cdots \pi_\epsilon\pi_{\epsilon+1} \cdots \pi_p; \alpha)$ as $r(, [\tau'] [\tau'']; [\pi'] [\pi'']; \alpha)$, where $[\tau'] = \tau_1, \cdots, \tau_\epsilon$, $[\tau''] = \tau_{\epsilon+1}, \cdots, \tau_p$, $[\pi'] = \pi_1, \cdots, \pi_\epsilon$, $[\pi''] = \pi_{\epsilon+1}, \cdots, \pi_p$.

We define a *δ-matrix* on $\tau_1, \cdots, \tau_n$ to be the matrix $(\delta_{\tau_1 \cdots \tau_n})$ for which $\delta_{\tau_1^* \cdots \tau_n^*} = 1$ when $\tau_1^* = \cdots = \tau_n^*$, and $\delta_{\tau_1^* \cdots \tau_n^*} = 0$ for all other values of $\tau_1, \cdots, \tau_n$. Evidently Kronecker's *delta* is the special case in which $n = 2$.

3. **The ranks of multiple composites.** We shall denote a matrix A by $(a_{\rho\tau_1 \cdots \tau_\epsilon\gamma_1 \cdots \gamma_\xi})$. A composite of a set of matrices $A_1, \cdots, A_n$ combined in any way will be denoted by $B = (b_{\rho'\tau_1' \cdots \tau_1\alpha\tau_2' \cdots \tau_2\beta \cdots \tau_\epsilon' \cdots \tau_\epsilon\sigma\lambda_1 \cdots \lambda_\nu})$,

[1] Presented to the American Mathematical Society, April, 1934.

where we assume that $\tau_1 = \tau_1' = \cdots = \tau_1^{\alpha}$, $\tau_2 = \tau_2' = \cdots = \tau_2^{\beta}$, $\cdots$, $\tau_\epsilon = \tau_\epsilon' = \cdots = \tau_\epsilon^{\sigma}$. The multiple composite of A and B obtained by combining τ_1 with τ_1', $\cdots$, τ_1^{α} not summing (the result of combining these partitions will be denoted by τ_{11}), $\cdots$, τ_ϵ with τ_ϵ', $\cdots$, τ_ϵ^{σ} not summing (denoted by $\tau_{\epsilon 1}$), and finally ρ with ρ' (assuming that $\rho = \rho'$) summed will be denoted by $C = (c_{\tau_{11} \cdots \tau_{\epsilon 1} \gamma_1 \cdots \gamma_\xi \lambda_1 \cdots \lambda_\nu})$. Evidently C is the most general multiple composite of a set of matrices.

Let $[\delta_1]$, $[\delta_2]$ denote sets of partitions of C where these sets contain partitions simply covered by τ's alone, partitions simply covered by and each containing both τ's and γ's or τ's and λ's, and partitions simply covered by and each containing τ's, γ's, and λ's. Let the γ's in $[\delta_1]$ be denoted by $\gamma_1, \cdots, \gamma_r$ and the γ's in $[\delta_2]$ by $\gamma_{r+1}, \cdots, \gamma_{r+s}$. Let the λ's in $[\delta_1]$ and $[\delta_2]$ be denoted by $\lambda_1, \cdots, \lambda_t$ and $\lambda_{t+1}, \cdots, \lambda_{t+u}$ respectively.

Let $[\rho_1]$, $[\rho_2]$ denote sets of partitions of C where these sets contain partitions simply covered by λ's, and partitions simply covered by and each containing both γ's and λ's. Let the γ's in $[\rho_1]$ be denoted by $\gamma_{r+s+1}, \cdots, \gamma_{r+s+p}$, and the γ's in $[\rho_2]$ by $\gamma_{r+s+p+1}, \cdots, \gamma_{r+s+p+w}$.

Let $[\gamma_1]$, $[\gamma_2]$ denote sets of partitions containing only γ's.

We can now state

THEOREM 1. *Let ω_1 denote the partition simply covered by $\gamma_1, \cdots, \gamma_{r+s+p+w}, \rho, \tau_1, \cdots, \tau_\epsilon$. Let π be the smallest of the orders of the partitions contained in $[\delta_1]$, $[\rho_1]$. Then:*

1. *The rank $r([\delta_1]\,[\rho_1]\,[\gamma_1], [\delta_2]\,[\rho_2]\,[\gamma_2])$ of C, where $[\gamma_1]$ is not vacuous, is less than or equal to the rank $r([\gamma_1]\omega_1, [\gamma_2])$ or $r([\gamma_1], \omega_1[\gamma_2])$ of A according as $[\gamma_1]$ contains an odd or even number of partitions of A.*

2. *The rank $r([\delta_1]\,[\rho_1], [\delta_2]\,[\rho_2]\,[\gamma_2])$ of C is less than or equal to the rank $r(, [\gamma_2]\omega_1; [\pi]\pi; \pi)$ of A, where $[\pi]$ is a set of μ integers if $[\gamma_2]$ contains μ partitions.*

Let the δ-matrix on partitions $\overset{*}{\tau_1}, \tau_1, \tau_1', \cdots, \tau_1^{\alpha}$ be denoted by $I_1, \cdots$, the δ-matrix on partitions $\overset{*}{\tau_\epsilon}, \tau_\epsilon, \tau_\epsilon', \cdots, \tau_\epsilon^{\sigma}$ by I_ϵ. Let the δ-matrix on ρ, ρ' be denoted by I_ρ. Let further I_q' be a δ-matrix on γ_q, γ_q' for $q = 1, \cdots, r+s+p+w$. Let finally I_q'' be a δ-matrix on λ_q, λ_q' for $q = 1, \cdots, t+u$.

Let $[\delta_1']$, $[\delta_2']$ be the sets of partitions obtained from $[\delta_1]$, $[\delta_2]$ by replacing $\tau_{11}, \cdots, \tau_{\epsilon 1}$ by $\overset{*}{\tau_1}, \cdots, \overset{*}{\tau_\epsilon}$; γ_q by γ_q', $q = 1, \cdots, r+s$; and λ_q by λ_q' for $q = 1, \cdots, t+u$.

Now the matrix C is equal to the matrix

$$C' = I_1 \times \cdots \times I_\epsilon \times I_\rho \times I_1' \times \cdots \times I_{r+s}' \times I_1'' \times \cdots \times I_{t+u}'' \,|\omega|\, A \times B,$$

where $\omega = \tau_1\tau_1' \cdots \tau_1^{\alpha} \cdots \tau_\epsilon\tau_\epsilon' \cdots \tau_\epsilon^{\sigma}\rho\rho'\gamma_1 \cdots \gamma_{r+s}\lambda_1 \cdots \lambda_{t+u}$. The rank $r([\delta_1]\,[\rho_1]\,[\gamma_1], [\delta_2]\,[\rho_2]\,[\gamma_2])$ of C is equal to the rank $r([\delta_1']\,[\rho_1]\,[\gamma_1], [\delta_2']\,[\rho_2]\,[\gamma_2])$ of C'. By Theorem 22 of the paper of which this is a supplement, if $[\rho_1]$, $[\gamma_1]$ are not both vacuous, the rank $r([\delta_1']\,[\rho_1]\,[\gamma_1], [\delta_2']\,[\rho_2]\,[\gamma_2])$ of C is less than or equal to the rank $r([\rho_1]\,[\gamma_1]\omega, [\rho_2]\,[\gamma_2])$ or $r([\rho_1]\,[\gamma_1], \omega[\rho_2]\,[\gamma_2])$ of $A \times B$

according as the sets $[\rho_1]$, $[\gamma_1]$ contain an odd or even number of partitions. If $[\rho_1]$, $[\gamma_1]$ are both vacuous we apply the corollary of Theorem 24. Further

$$(1) \qquad A \times B = A\,|\omega_1|\; B \times I'_1 \times \cdots \times I'_{r+s+p+w} \times I'_\rho \times I'''_1 \times \cdots \times I'''_\epsilon,$$

where I'''_s is a δ-matrix on τ_s, τ_s^{**} for $s = 1, \cdots, \epsilon$, and I'_ρ is a δ-matrix on ρ, ρ''. Again applying Theorem 22 and the corollary of Theorem 24 now to (1) Theorem 1 is proved.

For the non-signant ranks of C we prove

THEOREM 2. *Let ω_2 be the partition simply covered by $\gamma_1, \cdots, \gamma_{s+w}, \rho, \tau_1, \cdots, \tau_\epsilon$. The rank $r(, [\delta_2]\,[\rho_2]\,[\gamma_2];\, [\pi_1]\,[\pi_2]\,[\pi_3];\, \alpha)$ of C is less than or equal to the rank $r(, [\gamma_2]\omega_2;\, [\pi_3]\alpha;\, \alpha)$ of A.*

We now write C in the form

$$C'' = I_1 \times \cdots \times I_\epsilon^* \times I_\rho \times I'_1 \times \cdots \times I'_s \times I''_1 \times \cdots \times I''_u \quad |\omega_3| \quad A \times B,$$

where $\omega_3 = \tau_1\tau'_1 \cdots \tau_1^\alpha \cdots \tau_\epsilon\tau'_\epsilon \cdots \tau_\epsilon^\sigma\rho\rho'\gamma_1 \cdots \gamma_s\lambda_1 \cdots \lambda_u$. The rank $r(, [\delta_2]\,[\rho_2]\,[\gamma_2];\, [\pi_1]\,[\pi_2]\,[\pi_3];\, \alpha)$ of C is equal to the rank $r(, [\delta'_2]\,[\rho_2]\,[\gamma_2];\, [\pi_1]\,[\pi_2]\,[\pi_3];\, \alpha)$ of C''. By the corollary of Theorem 24 mentioned above $r(, [\delta'_2]\,[\rho_2]\,[\gamma_2];\, [\pi_1]\,[\pi_2]\,[\pi_3];\, \alpha)$ of C'' is less than or equal to the rank $r(, [\rho_2]\,[\gamma_2]\omega_3;\, [\pi_2]\,[\pi_3]\alpha;\, \alpha)$ of $A \times B$. Further

$$(2) \qquad A \times B = A\,|\omega_2|\; B \times I'_1 \times \cdots \times I'_{s+w} \times I'_\rho \times I'''_1 \times \cdots \times I'''_\epsilon.$$

We now apply the corollary of Theorem 24 to (2).

In Theorems 1 and 2 we have obtained relations between all of the space ranks of the matrix C and the space ranks of A, whence the theory is complete.

4. **Non-singular case.** We assume that the matrix $B = (b_{\rho\tau_1 \cdots \tau_\epsilon\lambda})$ is non-singular on $\rho\tau_1 \cdots \tau_\epsilon, \lambda$. Let the matrix $B^{-1} = (B_{\lambda,\ \rho'\tau_1' \cdots \tau_\epsilon'})$, $\rho = \rho'$, $\tau_1 = \tau'_1, \cdots, \tau_\epsilon = \tau'_\epsilon$, be the inverse of the display $(b_{\omega_4\lambda})$ where $\omega_4 = \rho\tau_1 \cdots \tau_\epsilon$. Then

$$(b_{\rho\tau_1 \cdots \tau_\epsilon\lambda}B_{\lambda\rho'\tau_1' \cdots \tau_\epsilon'}) = (\delta_{\rho\tau_1 \cdots \tau_\epsilon\rho'\tau_1' \cdots \tau_\epsilon'}),$$

where $\delta = (\delta_{\rho\tau_1 \cdots \tau_\epsilon\rho'\tau_1' \cdots \tau_\epsilon'})$ is a δ-matrix on the partitions $\rho\tau_1 \cdots \tau_\epsilon$, $\rho'\tau'_1 \cdots \tau'_\epsilon$. Now

$$(3) \qquad \delta = (\delta_{\rho\rho'}) \times (\delta_{\tau\tau'}) \times \cdots \times (\delta_{\tau_\epsilon\tau_\epsilon'}),$$

where $(\delta_{\rho\rho'})$, $(\delta_{\tau_1\tau_1'})$, $\cdots$, $(\delta_{\tau_\epsilon\tau_\epsilon'})$ are δ-matrices on ρ, ρ'; τ_1, τ'_1; $\cdots$; $\tau_\epsilon, \tau'_\epsilon$ respectively.

Let

$$(4) \qquad C = (a_{\rho\tau_1 \cdots \tau_\epsilon\gamma_1 \cdots \gamma_\xi}b_{\rho\tau_1 \cdots \tau_\epsilon\lambda}) \qquad (\tau_1, \cdots, \tau_\epsilon \text{ not summed}).$$

Now $C\,|\lambda|\,B^{-1} = (a_{\rho\tau_1 \cdots \tau_\epsilon\gamma_1 \cdots \gamma_\xi}\, b_{\rho\tau_1 \cdots \tau_\epsilon\lambda}B_{\lambda\rho'\tau_1' \cdots \tau_\epsilon'}) = (a_{\rho\tau_1 \cdots \tau_\epsilon\gamma_1 \cdots \gamma_\xi}\, \delta_{\rho\tau_1 \cdots \tau_\epsilon\rho'\tau_1' \cdots \tau_\epsilon'})$, which by (3) gives

$$(5) \qquad C\,|\lambda|\,B^{-1} = (a_{\rho'\tau_1 \cdots \tau_\epsilon\gamma_1 \cdots \gamma_\xi}\delta_{\tau_1\tau_1'} \cdots \delta_{\tau_\epsilon\tau_\epsilon'}) = A'.$$

Now the ranks of A' which involve the partition $\chi = \tau_1\tau'_1 \cdots \tau_\epsilon\tau'_\epsilon$ are equal to the corresponding ranks of A with χ replaced by $\tau_1 \cdots \tau_\epsilon$. Applying Theorems 1 and 2 to (4) and (5), we obtain

THEOREM 3. *Let* $(b_{\rho\tau_1 \cdots \tau_\epsilon\lambda})$ *be non-singular on* $\rho\tau_1 \cdots \tau_\epsilon, \lambda$. *The ranks of* $(a_{\rho\tau_1 \cdots \tau_\epsilon\gamma_1 \cdots \gamma_\xi} b_{\rho\tau_1 \cdots \tau_\epsilon\lambda})$, $\tau_1, \cdots, \tau_\epsilon$ *not summed, which involve the partition* $\tau_1 \cdots \tau_\epsilon\lambda$, *are equal to the corresponding ranks of* $(a_{\rho\tau_1 \cdots \tau_\epsilon\gamma_1 \cdots \gamma_\xi})$ *where* λ *has been replaced by* ρ.

CASE SCHOOL OF APPLIED SCIENCES.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

ON MECHANICAL QUADRATURES FORMULAE INVOLVING THE CLASSICAL ORTHOGONAL POLYNOMIALS[1]

BY C. WINSTON

(Received December 8, 1932)

Introduction. Let $\psi(x)$ be a function bounded and non-decreasing, with infinitely many points of increase, over an interval (a, b), finite or infinite, and such that all "moments" $\alpha_n = \int_a^b x^n d\psi(x)$ exist $(n = 0, 1, \cdots)$, with $\alpha_0 > 0$. It gives rise, as is known, to an infinite sequence of orthogonal and normal Tchebycheff polynomials

$$\varphi_n(x; a, b; d\psi(x)) \equiv \varphi_n(x; d\psi) \equiv \varphi_n(x) \equiv a_n(d\psi)\phi_n(x; d\psi),$$

(1)

$$\phi_n(x; d\psi) \equiv \phi_n(x) \equiv x^n - S_n x^{n-1} + \cdots \qquad (a_n(d\psi) \equiv a_n > 0; n = 0, 1, \cdots).$$

The sequences $\{\varphi_n(x; d\psi)\}$, $\{\phi_n(x; d\psi)\}$ are uniquely determined by the respective relations

$$\int_a^b \varphi_n(x)\varphi_m(x)d\psi(x) = \delta_{nm} \qquad (n, m = 0, 1, \cdots),$$

(2)

$$\int_a^b \phi_n(x)G_{n-1}(x)d\psi(x) = 0 \qquad (n = 1, 2, \cdots),$$

where hereafter $G_s(x)$ generally denotes an arbitrary polynomial of degree $\leqq s$.

A special case of (1) is derived by letting $d\psi(x) = p(x)dx$, $p(x) \geqq 0$ in (a, b). Here we use the notations

$$\varphi_n(x; a, b; p(x)) \equiv \varphi_n(x; p), \quad \phi_n(x; a, b; p(x)) \equiv \phi_n(x; p), \quad a_n(p), \quad S_n(p), \cdots.$$

One of the most important properties of the polynomials (1) is that the zeros $\{x_{in,}\}$ of $\phi_n(x)$ are all real, distinct and lie in (a, b):

$$a < x_{1,n} < x_{2,n} < \cdots < x_{n,n} < b \qquad (x_{i,n}(a, b; d\psi(x)) \equiv x_{i,n}(d\psi) \equiv x_{i,n}).$$

Making use of the Lagrange interpolation polynomial coinciding in value with a given function $f(x)$ at the above points $\{x_{i,n}\}$, we construct a formula of mechanical quadratures of Gauss' type:

[1] The author wishes to express his gratitude to Professor J. Shohat for his valuable suggestions and very helpful criticism.

$$(3)\qquad \int_a^b f(x)\,d\psi(x) = \sum_{i=1}^{n} H_{i,\,n} f(x_{i,\,n}) + R_n(f)$$

$$\left(H_{i,\,n} = \int_a^b \frac{\phi_n(x)\,d\psi(x)}{(x - x_{i,\,n})\,\phi_n'(x_{i,\,n})}\right),$$

with the remarkable property (a ready consequence of (2)):

$$\int_a^b G_{2n-1}(x)\,d\psi(x) = \sum_{i=1}^{n} H_{i,\,n} G_{2n-1}(x_{i,\,n}), \qquad \text{i.e.,} \qquad R_n(G_{2n-1}(x)) \equiv 0.$$

We call (3) of Gauss' type since Gauss [1][2] was first to consider (3) for the simplest case: $\psi(x) \equiv x$, (a, b) finite, $\varphi_n(x)$ being the Legendre polynomial.

The sequence (1) is said to be "symmetric," if

$$(a, b) \equiv (-h, h),\ \psi(x) \equiv -\psi(-x) \text{ (or } p(x) \equiv p(-x)).$$

It follows readily from (2, 3) that here, for any n,

$$(4)\qquad \begin{aligned} &\phi_n(x; d\psi) \equiv (-1)^n \phi_n(-x; d\psi); \qquad x_{i,\,n} + x_{n+1-i,\,n} = 0; \\ &\qquad\qquad H_{i,\,n} = H_{n+1-i,\,n}; \qquad (i = 1, 2, \cdots, n) \\ &\phi_{2n}(x; -h, h; d\psi) \equiv \phi_n(x^2; 0, h^2; d\psi_1); \qquad \phi_{2n+1}(x; -h, h; d\psi) \\ &\qquad\qquad \equiv x\phi_n(x^2; 0, h^2; d\psi_2); \\ &(d\psi_1(x) \equiv 2d\psi(\sqrt{x}), \qquad d\psi_2(x) \equiv x d\psi_1(x)). \end{aligned}$$

From (3, 4) we get at once:

$$(5)\qquad \begin{aligned} &x_{i,\,2n}(-h, h; d\psi) = -x_{2n+1-i,\,2n}(-h, h; d\psi) = x_{i-n,\,n}(0, h^2; d\psi_1), \\ &\qquad\qquad (i = n + 1, n + 2, \cdots, 2n) \\ &H_{i,\,2n}(-h, h; d\psi) = H_{2n+1-i,\,2n}(-h, h; d\psi) = \tfrac{1}{2} H_{i-n,\,n}(0, h^2; d\psi_1), \\ &x_{i,\,2n+1}(-h, h; d\psi) = -x_{2n+2-i,\,2n+1}(-h, h; d\psi) = x_{i-n-i,\,n}(0, h^2; d\psi_2), \\ &\qquad\qquad (i = n + 2, \cdots, 2n + 1) \\ &H_{i,\,2n+1}(-h, h; d\psi) = H_{2n+2-i,\,2n+1}(-h, h; d\psi) = \frac{H_{i-n-1,\,n}}{2x_{i-n-1,\,n}}\,(0, h^2; d\psi_2). \end{aligned}$$

In connection with formula (3) there naturally arises the question of convergence: for what functions $f(x)$ is $\lim_{n\to\infty} R_n(f) = 0$? For (a, b) finite and $d\psi(x) = p(x)dx$ this was almost completely answered by Stieltjes [2]. In recent years the case of (a, b) infinite was treated by several writers [3, 4].

It is evident that the convergence of (3) is intimately related to two questions: (i) how do the coefficients $\{H_{i,\,n}\}$ behave for n very large? (ii) how are the zeros $\{x_{i,\,n}\}$ distributed over (a, b) for n very large?

[2] The numbers in [ ] refer to the bibliography at the end of the paper.

The object of this paper is a study of questions (i) and (ii) for the case of the so-called classical orthogonal polynomials—of Jacobi, Laguerre and Hermite:

$$\text{J: } d\psi(x) = (x-a)^{\alpha-1}(b-x)^{\beta-1}\,dx, \qquad (a, b) \text{ finite}, \qquad \alpha, \beta > 0.$$

(6)[3] $$\text{L: } d\psi(x) = x^{\alpha-1}e^{-x}\,dx, \qquad (a, b) \equiv (0, \infty), \qquad \alpha > 0.$$

$$\text{H: } d\psi(x) = e^{-x^2}\,dx, \qquad (a, b) \equiv (-\infty, \infty).$$

From (4, 5, 6) we note the following relations between Hermite and Laguerre polynomials:

$$\phi_{2n}(x; -\infty, \infty; e^{-x^2}) \equiv \phi_n(x^2; 0, \infty; x^{-\frac{1}{2}}e^{-x});$$

$$\phi_{2n+1}(x; -\infty, \infty; e^{-x^2}) \equiv x\phi_n(x^2; 0, \infty; x^{\frac{1}{2}}e^{-x});$$

$$x_{i,2n}(-\infty, \infty; e^{-x^2}) = -x_{2n+1-i,2n}(-\infty, \infty; e^{-x^2}) = x_{i-n,n}(0, \infty; x^{-\frac{1}{2}}e^{-x})$$
$$(i = n+1, \cdots, 2n).$$

(7) $$H_{i,2n}(-\infty, \infty; e^{-x^2}) = H_{2n+1-i,2n}(-\infty, \infty; e^{-x^2}) = \tfrac{1}{2}H_{i-n,n}(0, \infty; x^{-\frac{1}{2}}e^{-x}),$$

$$x_{i,2n+1}(-\infty, \infty; e^{-x^2}) = -x_{2n+2-i,2n+1}(-\infty, \infty; e^{-x^2}) = x_{i-n-1,n}(0, \infty; x^{\frac{1}{2}}e^{-x}),$$
$$(i = n+2, \cdots, 2n+1)$$

$$H_{i,2n+1}(-\infty, \infty; e^{-x^2}) = H_{2n+2-i,2n+1}(-\infty, \infty; e^{-x^2}) = \frac{H_{i-n-1,n}}{2x_{i-n-1,n}}(0, \infty; x^{\frac{1}{2}}e^{-x}).$$

It is generally known that all the coefficients $H_{i,n} \to 0$ as $n \to \infty$. Furthermore, Sonine [5], in the case of Hermite polynomials, was able to show, by a very ingenious analysis, that

$$H_{i,n} < \frac{\pi}{\sqrt{2n \mp 1}}, \qquad \left(\mp \text{ for } n \begin{matrix}\text{even}\\ \text{odd}\end{matrix}; i = 1, 2, \cdots, n\right).$$

Regarding the zeros $\{x_{i,n}\}$ we know they are everywhere dense in every finite part of (a, b), with $x_{1,n} \to a$, $x_{n,n} \to b$, $(n \to \infty)$. E. R. Neumann [6] has shown, for Laguerre polynomials proper, i.e., with $\alpha = 1$, by means of interesting geometrical considerations, that

$$x_{i,n} = C_{i,n}\frac{(i+1)^2}{n+1}; \qquad \tfrac{1}{4} < C_{i,n} < 4 \qquad (i = 1, 2, \cdots, n).$$

In what follows, regarding question (i), we extend Sonine's method so as not only to embrace all classical polynomials, but also to obtain both upper and lower bounds for the coefficients $\{H_{i,n}\}$, their respective order of magnitude, their order with respect to n, and, in some cases, their asymptotic expression for $n \to \infty$. Regarding question (ii), we extend Neumann's method so as to include the general case of Laguerre polynomials for any $\alpha(> 0)$, whence the case of Hermite polynomials follows at once (by (7)).

[3] J, L, H stand, hereafter, for Jacobi, Laguerre and Hermite polynomials respectively.

This paper also contains a generalization of the mechanical quadratures formula where some of the points, at which the value of the function $f(x)$ to be integrated is known, are fixed.

1. **Some properties of the classical orthogonal polynomials.** These will be needed in our discussion.

$$A(x)\phi_n''(x) + B(x)\phi_n'(x) + C_n\phi_n(x) = 0 \qquad (n > 0).$$

$$A(x) = hx^2 + kx + l, \qquad B(x) = tx + u. \qquad (h, k, l, t, u, C_n \text{ const.})$$

(8)
$$\phi_{n+1}(x) = (x - c_{n+1})\phi_n(x) - \lambda_{n+1}\phi_{n-1}(x). \qquad (n \geqq 1)$$

$$c_{n+1} = S_{n+1} - S_n; \qquad \lambda_{n+1} = \frac{a_{n-1}^2}{a_n^2}. \qquad (n > 0)$$

H: $A(x) \equiv 1$; $B(x) \equiv -2x$; $C_n = 2n$; $\lambda_{n+1} = n/2$.

L: $A(x) \equiv x$; $B(x) \equiv \alpha - x$; $C_n = n$; $\lambda_{n+1} = n(n + \alpha - 1)$.

(9) J: $A(x) \equiv 1 - x^2$; $B(x) \equiv \alpha - \beta - (\alpha + \beta)x$; $C_n = n(n + \alpha + \beta - 1)$;

$$\lambda_{n+1} = \frac{4n(n + \alpha + \beta - 2)(n + \alpha - 1)(n + \beta - 1)}{(2n + \alpha + \beta - 1)(2n + \alpha + \beta - 2)^2(2n + \alpha + \beta - 3)}.$$

(Hereafter for J: $(a, b) \equiv (-1, 1)$).

The following relation is fundamental for our discussion:

(10)
$$\phi_n'(x; p) \equiv n\phi_{n-1}(x; Ap) \qquad (n \geqq 1; A \equiv A(x) \text{ in (8)}).$$

Hereafter we shall use the abbreviated notations:

$$\phi_n(x; p) \equiv \phi_n(x), \qquad \phi_n(x; Ap) \equiv \phi_n^*(x), \qquad a_n(p) \equiv a_n, \qquad a_n(Ap) \equiv a_n^*, \cdots.$$

Other necessary formulae are:

(11)
$$H_{i,n}{}^{[7]} = \frac{1}{a_{n-1}^2\phi_{n-1}(x_{i,n})\phi_n'(x_{i,n})} \qquad (i = 1, 2, \cdots, n).$$

(12) H:
$$\begin{cases} \displaystyle\sum_{i=1}^{n} H_{i,n} = \sqrt{\pi}, \qquad \phi_{2n}(0) = (-1)^n \frac{(2n)!}{2^{2n}\, n!}, \\ \phi_{2n+1}(0) = 0, \qquad a_n^2 = \dfrac{2^n}{n!\sqrt{\pi}}. \\ \phi_{n-1}^{(x)}{}^{[8]} = \dfrac{(n-1)!\, e^{\frac{n}{2} + \frac{x^2}{2}}}{2^{\frac{n-1}{2}}\sqrt{\pi}\, n^{\frac{n}{2}} \sqrt[4]{1 - \frac{x^2}{2n}}} \left\{ \sin\left(\frac{n\pi}{2} - x\sqrt{2n}\right) + 0\left(\frac{1}{\sqrt{n}}\right) \right\} \\ \qquad\qquad x^2 < 2n\text{-}n^{\frac{1}{2}+\epsilon}, \epsilon > 0 \end{cases}$$

$$
(13)\ \mathrm{L}:\begin{cases}
\displaystyle\sum_{i=1}^{n} H_{i,n} = \Gamma(\alpha); \quad \phi_n^2(0) = \left[\frac{\Gamma(n+\alpha)}{\Gamma(\alpha)}\right]^2; \quad a_n^2 = \frac{1}{\Gamma(n+1)\Gamma(n+\alpha)}; \\
\phi_n(x)^{[9]} = \dfrac{(-1)^n}{\sqrt{\pi}}\, n!\, x^{\frac{1}{4}-\frac{\alpha}{2}} e^{\frac{x}{2}} n^{\frac{\alpha}{2}-\frac{3}{4}} \left\{\cos\left(2\sqrt{xn} + \frac{\pi}{4} - \frac{\pi\alpha}{2}\right) + \frac{x^{\frac{3}{2}}}{\sqrt{n}}\right\} \\
\qquad\qquad 0 < a < x < kn^{\frac{1}{3}}\ (a, k \text{ const.})
\end{cases}
$$

$$
(14)\ \mathrm{J}:\begin{cases}
\phi_n(1) = \dfrac{2^n \Gamma(n+\beta)\Gamma(n+\alpha+\beta-1)}{\Gamma(\beta)\Gamma(2n+\alpha+\beta-1)}; \\
\phi_n(-1) = \dfrac{(-1)^n 2^n \Gamma(n+a)\Gamma(n+\alpha+\beta-1)}{\Gamma(\alpha)\Gamma(2n+\alpha+\beta-1)}; \\
\phi_{2n}(0) = \dfrac{(-1)^n (2n)!}{2^n n!\,(2\beta+4n-3)(2\beta+4n-5)\cdots(2\beta+2n-1)} \quad (\alpha=\beta); \\
\varphi_n(x)^{[10]} = (-1)^n \sqrt{\dfrac{(2n+\alpha+\beta-1)\,\Gamma(n+\alpha+\beta-1)\,\Gamma^2(n+\frac{1}{2})}{\pi 2^{\alpha+\beta-1}\,\Gamma(n+1)\,\Gamma(n+\alpha)\,\Gamma(n+\beta)}} \\
\quad \left\{(\sin\theta)^{\frac{1}{2}-\alpha}(\cos\theta)^{\frac{1}{2}-\beta}\left[1 - \dfrac{(2\alpha-1)(2\beta-1)}{4(2n-1)}\right]\right. \\
\quad \times \cos[(2n+\alpha+\beta-1)\theta - \pi/4\,(2\alpha-1)] \\
\quad - \sin[(2n+\alpha+\beta-1)\theta - \pi/4\,(2\alpha-1)] \\
\quad \left.\times\left[\dfrac{(2\alpha-1)(2\alpha-3)}{8(2n-1)}\cot\theta - \dfrac{(2\beta-1)(2\beta-3)}{8(2n-1)}\tan\theta\right] + 0\left(\dfrac{1}{n^2}\right)\right\} \\
\qquad\qquad \left(\sin^2\theta = \dfrac{1+x}{2}\right).
\end{cases}
$$

2. **The behavior of the coefficients $H_{i,n}$.** We transform (11) by means of (10) as follows. Write

$$A(x)\,\varphi_n^*(x) = \sum_{i=0}^{n+2} B_i\,\varphi_i(x) \qquad (A(x) \text{ given in (8)});$$

$$
(15)\quad B_i = 0\ (0 \leqq i \leqq n-1); \qquad B_n = \frac{a_n}{a_n^*}; \qquad B_{n+1} = \frac{a_{n+1}}{a_n^*}\,(S_{n+1}^* - S_{n+1});
$$

$$B_{n+2} = \frac{h a_n^*}{a_{n+2}} \text{ (by (2, 8, 9))}.$$

Substituting these values in (15) and using (8, 10), we get:

$$
(16)\quad \phi_{n-1}(x_{i,n}) = \frac{A(x_{i,n})\,a_{n-1}^{*2}\,\phi_n'(x_{i,n})}{n(a_{n-1}^2 - h a_{n-1}^{*2}\,\lambda_{n+1})}; \qquad H_{i,n} = \frac{n(a_{n-1}^2 - h a_{n-1}^{*2}\,\lambda_{n+1})}{a_{n-1}^2\,a_{n-1}^{*2}\,A(x_{i,n})\,\phi_n'(x_{i,n})}
$$

$$(i = 1, 2, \cdots, n).$$

Following Sonine, we write

$$H_{i,n} = \frac{n(a_{n-1}^2 - ha_{n-1}^{*2}\lambda_{n+1})}{a_{n-1}^2 a_{n-1}^{*2}(A(x_{i,n})\phi_n'(x_{i,n}) + \sigma(x_{i,n})\phi_n(x_{i,n}))} \tag{17}$$

$$\equiv \frac{n(a_{n-1}^2 - ha_{n-1}^{*2}\lambda_{n+1})}{a_{n-1}^2 a_{n-1}^{*2}\gamma(x_{i,n})},$$

$$\gamma(x) = A(x)\phi_n'^2(x) + \sigma(x)\phi_n(x), \tag{18}$$

where $\sigma(x)$ is a certain function of x, finite at $x = x_{i,n}$ for all n and $i = 1, 2, \cdots, n$, and so chosen that $\gamma'(x)$ *be divisible by* $\phi_n'^2(x)$. Differentiate (18) and make use of (8):

$$\gamma'(x) = \phi_n'^2(x)[A'(x) - 2B(x)] - 2C_n\phi_n(x)\phi_n'(x) + [\sigma(x)\phi_n(x)]'.$$

Consequently, $\gamma'(x)$ will certainly be divisible by $\phi_n'^2(x)$ if

$$[\sigma(x)\phi_n(x)]' = 2C_n\phi_n(x)\phi_n'(x), \qquad \sigma(x) = C_n\phi_n(x) + \frac{k}{\phi_n(x)} \qquad (k = \text{const.}).$$

Here k must be zero ($\sigma(x_{i,n})$ being finite) so that finally

$$\gamma(x) = A(x)\phi_n'^2(x) + C_n\phi_n^2(x). \tag{19}$$

We have further, in view of (8),

$$\gamma'(x) = 2(h-t)\phi_n'^2(x)\left[x - \frac{2u-k}{2(h-t)}\right] \qquad (h \neq t), \tag{20}$$

$$\gamma'(x) = \phi_n'^2(x)(k - 2u) \qquad (h = t) \tag{21}$$

(20, 21) lead directly to the following conclusions.

(i) $h \neq t$. $\gamma(x)$ is minimum or maximum at

$$x = x_0 = \frac{2u-k}{2(h-t)}$$

according as $h > t$ or $h < t$, and $\gamma\left(\frac{2u-k}{2(h-t)}\right)$ gives correspondingly an upper or lower bound for $H_{i,n}$.

(ii) $h = t$. This case is possible only for Jacobi polynomials with $\alpha + \beta = 1$ (see (9)). For $k > 2u$, $k < 2u$, $k = 2u$, $\gamma(x)$ is correspondingly monoton increasing, monoton decreasing, or constant. Hence,

$$k > 2u\colon H_{1,n} > H_{2,n} > \cdots > H_{n,n}.$$

$$k < 2u\colon H_{1,n} < H_{2,n} < \cdots < H_{n,n}. \tag{22}$$

$$k = 2u\colon H_{1,n} = H_{2,n} = \cdots = H_{n,n} = \frac{1}{n}\int_{-1}^{+1}\frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{n}.$$

The constancy of $H_{i,n}(i = 1, 2, \cdots, n)$ is possible for trigonometric polynomials only ($\alpha = \beta = \frac{1}{2}$), as is seen by (9, 22).

We now apply (19–21) to the polynomials J, L, H separately, making use of (12–14).

HERMITE POLYNOMIALS. Here

$$(23)\quad \gamma(x) = \phi_n'^2(x) + 2n\phi_n^2(x), \quad h > t, \quad x_0 = \frac{2u-k}{2(h-t)} = 0, \quad \gamma(x_0) \text{ is minimum},$$

$$H_{i,n} \leqq \frac{n}{a_{n-1}^2(\phi_n'^2(0) + 2n\phi_n^2(0))} \qquad (i = 1, 2, \cdots, n).$$

$$H_{i,2m} = H_{2m+1-i,2m} \leqq \frac{(m!)^2 2^{2m}\sqrt{\pi}}{2m(2m!)} = \frac{\pi\Gamma(m)}{2\Gamma(m+\frac{1}{2})} < \frac{\pi}{\sqrt{4m-1}} \quad (i = 1, 2, \cdots, 2m),$$

(24)[4]

$$H_{i,2m+1} = H_{2m+2-i,2m+i} \leqq \frac{(m!)^2 2^{2m}\sqrt{\pi}}{(2m+1)!} = \frac{\pi\Gamma(m+1)}{2\Gamma(m+\frac{3}{2})} < \frac{\pi}{\sqrt{4m+3}} \quad (i = 1, 2, \cdots, 2m+1).$$

Formulae (24) represent Sonine's results. The above considerations give even more:

$$H_{1,2m} < H_{2,2m} < \cdots < H_{m,2m} = H_{m+1,2m} > H_{m+2,2m} > \cdots > H_{2m,2m}$$

(since $x_{1,2m} < x_{2,2m} < \cdots < x_{m,2m} < 0 < x_{m+1,2m} < \cdots < x_{2m,2m}$). For $n = 2m + 1$ we have similarly:

$$H_{1,2m+1} < H_{2,2m+1} < \cdots < H_{m+1,2m+1} > \cdots > H_{2m+1,2m+1}.$$

$\gamma(x)$ being here an even function of x, it follows by (20, 23) that

$$(25)\quad s < |x_{i,n}| < t \text{ implies } \gamma(s) < \gamma(x_{i,n}) < \gamma(t); \quad \frac{n}{a_{n-1}^2\gamma(t)} < H_{i,n} < \frac{n}{a_{n-1}^2\gamma(s)};$$

$$(26)\quad \frac{\pi}{\sqrt{2n}\, e^{t^2}(1 + O(1/\sqrt{n}))} < H_{i,n} < \frac{\pi}{\sqrt{2n}\, e^{s^2}(1 + O(1/\sqrt{n}))}$$

$$(0 \leqq s < |x_{i,n}| < t < \sqrt{2n - n^{1+\epsilon}}\ \epsilon > 0).$$

It will be shown later (64, 65) that

$$\frac{i-m}{2\sqrt{m+1}} < -x_{2m+1-i,2m} = x_{i,2m} < \frac{2(i-m+1)}{\sqrt{m+1}}$$

$$(i = m+1, m+2, \cdots, 2m),$$

$$(27)\quad \frac{i-m}{2\sqrt{m+1}} < -x_{2m+2-i,2m+1} = x_{i,2m+1} < \frac{2(i-m+1)}{\sqrt{m+1}}$$

$$(i = m+2, \cdots, 2m+1).$$

---

[4] We use here the inequality [5]: $\dfrac{\Gamma(x)}{\Gamma(x+\frac{1}{2})} < \dfrac{2}{\sqrt{4x-1}}$ $(x > \frac{1}{4})$.

Hence, by (26),

$$\frac{\pi}{2\sqrt{m}\,e^{\frac{4(i-m+1)^2}{m+1}}(1+O(1/\sqrt{m}))} < H_{2m+1-i,\,2m}$$

$$= H_{i,\,2m} < \frac{\pi}{2\sqrt{m}\,e^{\frac{(i-m)^2}{4(m+1)}}(1+O(1/\sqrt{m}))} \qquad (i = m+1, \cdots, 2m),$$

(28)

$$\frac{\pi}{2\sqrt{m}\,e^{\frac{4(i-m+1)^2}{m+1}}(1+O(1/\sqrt{m}))} < H_{2m+2-i,\,2m+1}$$

$$= H_{i,\,2m+1} < \frac{\pi}{2\sqrt{m}\,e^{\frac{(i-m)^2}{4(m+1)}}(1+O(1/\sqrt{m}))} \qquad (i = m+2, \cdots, 2m+1).$$

From (28) we get an interesting asymptotic expression for some of the $H_{i,\,n}$:

(29) $$\lim_{m\to\infty} \sqrt{m}\,H_{i,\,2m+\sigma} = \pi/2, \qquad \text{i.e.,} \qquad H_{2m+\sigma+1+i,\,2m+\sigma} = H_{i,\,2m+\sigma} = \frac{\pi/2 + o(1)}{\sqrt{m}}$$

$$(\sigma = 0, 1;\ m - m^{\frac{1}{2}-\epsilon} < i < m + m^{\frac{1}{2}-\epsilon},\ \epsilon > 0).$$

Furthermore, the inequalities (28) enable us, in each case where the behavior of $f(x)$ for large $|x|$ is known, to estimate how much the mechanical quadratures formula (3) is affected if we omit even infinitely many of the terms $H_{i,\,n}f(x_{i,\,n})$ corresponding to distant $x_{i,\,n}(\gtrless 0)$.

Illustration. (i) Consider the mechanical quadratures formula (3) for $n = 2m + \sigma$ ($\sigma = 0, 1$) in which $f(x)$ is such that

$$|f(x)| = O(|x|^l) \qquad (|x| > A \text{ which is sufficiently large}).$$

Then we can omit in (3) all terms $H_{i,\,2m+\sigma}f(x_{i,\,2m+\sigma})$ corresponding to

(30) $$i > m + \sqrt{km\log m}, \qquad i < m - \sqrt{km\log m}$$

$$(k > \text{max. } (0,\ 2\,(l+1))).$$

(k here and hereafter generally denotes a properly fixed positive quantity.) In fact, for all such i, $|x_{i,\,2m+\sigma}| \to \infty$ as $m \to \infty$ (by (27)). Hence, for m sufficiently large, all such $|x_{i,\,2m+\sigma}| > A$. On the other hand, since

$$|x_{i,\,2m+\sigma}| < \sqrt{4m + 2\sigma + 1}$$

(see (53, 61))

(31) $$|\Sigma H_{i,\,2m+\sigma} f(x_{i,\,2m+\sigma})| = O\left(\frac{1}{m^{\frac{k}{4}-\frac{l+1}{2}}}\right) \to 0 \qquad (m \to \infty\ ;\ i \text{ and } k \text{ as in (30)}).$$

We see that *the number of terms actually used in* (3) *is only an infinitely small part of the total number of terms therein.*

(ii) Similar considerations show that if

$$|f(x)| = O(e^{|x|^l}) \qquad (l < 1; |x| > A),$$

we can omit in (3), without impairing the convergence, all terms

$$H_{i,\,2m+\sigma} f(x_{i,\,2m+\sigma})$$

for which $i > m + m^{\frac{1}{2}+\epsilon}$, $i < m - m^{\frac{1}{2}+\epsilon}$, $l/2 \leqq \epsilon < \frac{1}{2}$. Here the number of terms which we can afford to omit in (3) again increases indefinitely with m, but is smaller than in the preceding case, which was to be expected, since $|f(x)|$ increases much more strongly at infinity.

LAGUERRE POLYNOMIALS. Here.

$$\gamma(x) = x\phi_n'^2(x) + n\phi_n^2(x), \quad h > t, \quad x_0 = \alpha - \tfrac{1}{2}, \quad \gamma(x_0) \text{ is minimum.} \tag{32}$$

$$H_{i,\,n} \leqq \frac{n}{a_{n-1}^{*2}((\alpha - \frac{1}{2})\,\phi_n'^2(\alpha - \frac{1}{2}) + n\phi_n^2(\alpha - \frac{1}{2}))} \qquad (i = 1, 2, \cdots, n). \tag{33}$$

(i) $\alpha \leqq \frac{1}{2}$.

$$H_{1,\,n} > H_{2,\,n} > \cdots > H_{n,\,n},$$

$$H_{i,\,n} \leqq \frac{1}{a_{n-1}^{*2}\phi_n^2(0)} = \frac{\pi\Gamma(n)}{\Gamma(n + \frac{1}{2})} < \frac{2\pi}{\sqrt{4n-1}} \tag{34}$$

$(i = 1, 2, \cdots, n)$(by (9, 13, 24)).

The last inequality could also have been derived from (7, 24).

(ii) $\alpha > \frac{1}{2}$. As i increases, $H_{i,\,n}$ increases for $0 < x_{i,\,n} \leqq \alpha - \frac{1}{2}$ and decreases for $\alpha - \frac{1}{2} \leqq x_{i,\,n} < \infty$. By (13, 33), we get:

$$H_{i,\,n} < \left(\frac{\alpha - \frac{1}{2}}{e}\right)^{\alpha-\frac{1}{2}} \frac{\pi}{\sqrt{n}\,(1 + O(1/\sqrt{n}))} \qquad (i = 1, 2, \cdots, n), \tag{35}$$

with the same order of magnitude, with respect to n, on the right-hand side as for $\alpha = \frac{1}{2}$. Since here $\gamma(0) > \gamma(\alpha - \frac{1}{2})$, we have, by (13, 33, 35),

$$\frac{\Gamma^2(\alpha)}{(n+\alpha)^\alpha\, e^{\frac{1}{12(n+\alpha)}}} < H_{i,\,n} < \frac{(\alpha - \frac{1}{2})^{\alpha-\frac{1}{2}}\,\pi}{e^{\alpha-\frac{1}{2}}\sqrt{n}\,(1 + O(1/\sqrt{n}))} \quad (0 < x_{i,\,n} < \alpha - \tfrac{1}{2};\ \alpha > \tfrac{1}{2}).$$

Furthermore,

$$\frac{n}{a_{n-1}^{*2}\gamma(t)} < H_{i,\,n} < \frac{n}{a_{n-1}^{*2}\gamma(s)} \qquad (0 < \alpha - \tfrac{1}{2} < s < x_{i,\,n} < t),$$

$$\frac{\pi t^{\alpha-\frac{1}{2}}}{e^t\sqrt{n}\,(1 + O(1/\sqrt{n}))} < H_{i,\,n} < \frac{\pi s^{\alpha-\frac{1}{2}}}{e^s\sqrt{n}\,(1 + O(1/\sqrt{n}))} \tag{36}$$

$(\alpha - \frac{1}{2} < s < x_{i,\,n} < t < n^{\frac{1}{2}-\epsilon};\ \epsilon > 0)$(by (13)).

It will be shown later (59) that

$$\frac{(i+2\sqrt{\alpha}-1)^2}{4(n+1)} < x_{i,n} < \frac{4(i+\alpha)^2}{n+\alpha} \qquad \left(i=1,2,\cdots,n; \quad \alpha > \frac{(1+\sqrt{2})^2}{16}\right).$$

Hence

$$(37)\qquad \frac{\pi\left[\frac{4(i+\alpha)^2}{n+\alpha}\right]^{\alpha-\frac{1}{2}}}{e^{\frac{4(i+\alpha)^2}{n+\alpha}}\sqrt{n}\left(1+O\left(\frac{1}{\sqrt{n}}\right)\right)} < H_{i,n} < \frac{\pi\left[\frac{(i+2\sqrt{\alpha}-1)^2}{4(n+1)}\right]^{\alpha-\frac{1}{2}}}{e^{\frac{(i+2\sqrt{\alpha}-1)^2}{4(n+1)}}\sqrt{n}\left(1+O\left(\frac{1}{\sqrt{n}}\right)\right)}$$

$$k' n^{\frac{1}{2}} < i < k n^{\frac{1}{2}} \ (k, k' \text{ const.}), \epsilon > 0).$$

If in (37) we let $\alpha = \frac{1}{2}$, we get the asymptotic expression

$$H_{i,n} = \frac{\pi + o(1)}{\sqrt{n}} \qquad (n \to \infty; \quad 0 < i < n^{\frac{1}{2}-\epsilon}, \quad \epsilon > 0).$$

This result could have been derived from (7, 29).

As in the case of Hermite polynomials, we may, under certain conditions, omit infinitely many terms, without impairing the convergence of (3).

Illustration. If $|f(x)| = O(x^l)$ for $x > A > 0$ sufficiently large, we can omit in (3) all terms corresponding to $i > \sqrt{kn \log n}$ ($k > \max.\ (0, 4l+2)$). We could also obtain results similar to those in (ii) for the Hermite case.

JACOBI POLYNOMIALS. Here

$$\gamma(x) = (1-x^2)\phi_n'^2(x) + n(n+\alpha+\beta-1)\phi_n^2(x), \qquad H_{i,n} = \frac{\delta_n}{\gamma(x_{i,n})}$$

$$(38)\qquad (i = 1, 2, \cdots, n)$$

$$\delta_n = \frac{n(a_{n-1}^2 + a_{n-1}^{*2}\lambda_{n+1})}{a_{n-1}^2 a_{n-1}^{*2}}.$$

(i) $h \neq t$. $\gamma(x)$ has an extremum at

$$x_0 = \frac{\alpha-\beta}{\alpha+\beta-1}.$$

Case 1: $h > t$. Possible if and only if $\alpha+\beta > 1$.

a) $\beta \leqq \frac{1}{2}$. Here $x_0 = \dfrac{\alpha-\beta}{\alpha+\beta-1} \geqq 1$ for all α, β under consideration.

$\gamma(x_0)$ is minimum; $\qquad H_{1,n} < H_{2,n} < \cdots < H_{n,n}$.

We have further, since $\dfrac{\delta_n}{\gamma(-1)} < H_{i,n} < \dfrac{\delta_n}{\gamma(1)}$:

$$\frac{2^{\alpha+\beta-1}\Gamma^2(\alpha)}{n^{2\alpha}(1+O(1/n))} < H_{i,n} < \frac{2^{\alpha+\beta-1}\Gamma^2(\beta)}{n^{2\beta}(1+O(1/n))} \qquad (i = 1, 2, \cdots, n).$$

b) $\alpha \leqq \frac{1}{2}$. We derive from a), by interchanging α and β:

$$H_{1,n} > H_{2,n} > \cdots > H_{n,n}; \qquad \frac{2^{\alpha+\beta-1}\,\Gamma^2(\beta)}{n^{2\beta}(1+O(1/n))} < H_{i,n} < \frac{2^{\alpha+\beta-1}\,\Gamma^2(\alpha)}{n^{2\alpha}(1+O(1/n))}$$

$$(i = 1, 2, \cdots, n).$$

c) $\alpha > \frac{1}{2}, \beta > \frac{1}{2}, \alpha > \beta$. Here, $0 < x_0 = \dfrac{\alpha-\beta}{\alpha+\beta-1} < 1$, and as i increases, $H_{i,n}$ increases for $-1 < x_{i,n} \leqq x_0$ and decreases for $x_0 \leqq x_{i,n} < 1$. Hence, (by (9, 14)),

$$\frac{\delta_n}{\gamma(-1)} < H_{i,n} < \frac{\delta_n}{\gamma(x_0)}; \qquad \frac{2^{\alpha+\beta-1}\,\Gamma^2(\alpha)}{n^{2\alpha}(1+O(1/n))} < H_{i,n} < \frac{\pi(1+x_0)^{2\alpha}(1-x_0)^{2\beta}}{2^{\alpha+\beta+1}\,n(1+O(1/n))}$$

$$(-1 < x_{i,n} \leqq x_0);$$

$$\frac{\delta_n}{\gamma(1)} < H_{i,n} < \frac{\delta_n}{\gamma(x_0)}; \qquad \frac{2^{\alpha+\beta-1}\,\Gamma^2(\beta)}{n^{2\beta}(1+O(1/n))} < H_{i,n} < \frac{\pi(1+x_0)^{2\alpha}(1-x_0)^{2\beta}}{2^{\alpha+\beta+1}\,n(1+O(1/n))}$$

$$(x_0 \leqq x_{i,n} < 1).$$

d) $\alpha > \frac{1}{2}, \beta > \frac{1}{2}, \beta > \alpha$. The results are the same as for c), x_0 here lying in $(-1, 0)$.

e) $\alpha = \beta > \frac{1}{2}$ (Symmetric case). Here $x_0 = 0$, $\gamma(0)$ is minimum;

$$H_{1,2m} < H_{2,2m} < \cdots < H_{m,2m} = H_{m+1,2m} > \cdots > H_{2m,2m}.$$

$$H_{1,2m} = H_{2m,2m} < \frac{2^{2(\beta-1)}\,\Gamma^2(\beta)}{m\,\Gamma(2\beta)}; \qquad H_{m,2m} = H_{m+1,2m} > \frac{2^{2(\beta-1)}\,\Gamma^2(\beta)}{m\,\Gamma(2\beta)}.$$

$$H_{1,2m+1} < H_{2,2m+1} < \cdots < H_{m+1,2m+1} > H_{m+2,2m+1} > \cdots > H_{2m+1,2m+1};$$

$$H_{1,2m+1} = H_{2m+1,2m+1} < \frac{2^{2\beta-1}\,\Gamma^2(\beta)}{(2m+1)\,\Gamma(2\beta)}; \qquad H_{m+1,2m+1} > \frac{2^{2\beta-1}\,\Gamma^2(\beta)}{(2m+1)\,\Gamma(2\beta)};$$

$$\frac{\Gamma^2(\beta)\,2^{2\beta-1}}{m^{2\beta}\,(1+O(1/m))} < H_{i,2m+\sigma} < \frac{\pi\,2^{2\beta-1}}{m(1+O(1/m))}$$

$$(\sigma = 0, 1;\ i = 1, 2, \cdots, 2m+\sigma).$$

Case 2. $h < t$. Possible if and only if $\alpha + \beta < 1$.

a) $\beta > \alpha$. $x_0 = \dfrac{\alpha-\beta}{\alpha+\beta-1} \geqq 1$, $\gamma(x_0)$ is maximum;

$$H_{1,n} > H_{2,n} > \cdots > H_{n,n}; \frac{2^{\alpha+\beta-1}\,\Gamma^2(\alpha)}{n^{2\alpha}\,(1+O(1/n))} > H_{i,n} > \frac{2^{\alpha+\beta-1}\,\Gamma^2(\beta)}{n^{2\beta}\,(1+O(1/n))}$$

$$(i = 1, 2, \cdots, n).$$

b) $\alpha > \beta$. $x_0 \leqq -1$, $\gamma(x_0)$ is maximum. The results concerning $H_{i,n}$ are the same as in a), with inequality signs reversed.

c) $\alpha = \beta$ (Symmetric Case). $x_0 = 0, \gamma(0)$ is maximum. We get the same results as in case 1 e), with inequality signs reversed throughout.

(ii) $h = t$. Possible if and only if $\alpha + \beta = 1$.

a) $k > 2u$ $(\beta > \alpha)$, b) $k < 2u$ $(\alpha > \beta)$, c) $k = 2u$ $(\alpha = \beta = \frac{1}{2})$.

In a) and b) we get the same results respectively as in cases 1 a) and b). c) gives the trigonometric case discussed above (see (22)).

3. **A mechanical quadrature formula with fixed points.** We now fix m distinct points $\xi_1, \xi_2, \cdots, \xi_k, \eta_{k+1}, \eta_{k+2}, \cdots, \eta_m$ outside the interval (a, b) at which the values of the function $f(x)$ to be integrated are known. Making use of the orthogonality property (2) we readily derive

THEOREM I. *In order to have*

$$\int_a^b G_{2n-m-1}(x)\, d\psi(x) = \sum_{i=1}^{n} H_{i,n}\, G_{2n-m-1}(x_{i,n}),$$

$$x_{1,n} < x_{2,n} < \cdots < x_{k,n} \leqq a < x_{m+1,n} < x_{m+2,n} < \cdots < x_{n,n} \leqq x_{k+1,n} < x_{k+2,n} < \cdots < x_{m,n}, \qquad x_{i,n} = \xi_i \ (i = 1, 2, \cdots, k), \quad x_{i,n} = \eta_i \ (i = k+1, \cdots, m), \tag{39}$$

where

$$H_{i,n} = \int_a^b \frac{\theta_n(x)\, d\psi(x)}{(x - x_{i,n})\, \theta_n'(x_{i,n})}, \qquad \theta_n(x) = \prod_{i=1}^{n} (x - x_{i,n}),$$

it is necessary and sufficient that $x_{i,n}$ $(i = m+1, \cdots, n)$ *be the zeros of the Tchebycheff polynomial* $\varphi_{n-m}(x; a, b; d\psi_1(x))$, *with*

$$d\psi_1(x) = \prod_{i=1}^{k} (x - \xi_i) \prod_{i=k+1}^{m} (\eta_i - x)\, d\psi(x).$$

We proceed to derive some relations for the $H_{i,n}$ corresponding to the fixed points. Consider first the interval (a, b), a being finite, with one point fixed at $\xi \leqq a$. Letting $d\psi_1(x) = (x - \xi)\, d\psi(x)$, we have, by the above theorem:

$$H_{i,n} = \int_a^b \frac{(x - \xi)\varphi_{n-1}(x; d\psi_1(x))\, d\psi(x)}{(x - x_{i,n})\, [(x - \xi)\, \varphi_{n-1}(x; d\psi_1(x))]'_{x = x_{i,n}}} \qquad (i = 1, 2, \cdots, n),$$

$$x_{1,n} = \xi \leqq a, \qquad \varphi_{n-1}(x_{i,n}; d\psi_1(x)) = 0 \qquad (i = 2, 3, \cdots, n),$$

$$H_{1,n} = H_\xi = \int_a^b \frac{\varphi_{n-1}^*(x)\, d\psi(x)}{\varphi_{n-1}^*(\xi)}, \qquad (\varphi_n^*(x) \equiv \varphi_n(x; d\psi_1(x)), \text{ with } a_n^*, \cdots).$$

Using the expansion

$$(x - \xi)\, \varphi_{n-1}^*(x) = \frac{a_{n-1}^*}{a_n\, \varphi_{n-1}(\xi)} \frac{\varphi_n(x)\varphi_{n-1}(\xi) - \varphi_n(\xi)\varphi_{n-1}(x)}{(x - \xi)}$$

and the Darboux formulae [10]

$$K_{n-1}(x;\xi;d\psi) \equiv K_{n-1}(x;\xi) \equiv \sum_{i=0}^{n-1} \varphi_i(x)\varphi_i(\xi) = \frac{a_{n-1}}{a_n} \frac{\varphi_n(x)\varphi_{n-1}(\xi) - \varphi_n(\xi)\varphi_{n-1}(x)}{x-\xi}$$

$$K_{n-1}(\xi;d\psi) \equiv K_{n-1}(\xi;\xi;d\psi) \equiv \sum_{i=0}^{n-1} \varphi_i^2(\xi;d\psi) = \cdots,$$

we get

$$H_\xi^{[u]} = \frac{a_{n-1}^*}{a_{n-1}\varphi_{n-1}(\xi)\,\varphi_{n-1}^*(\xi)} \int_a^b \sum_{i=0}^{n-1} \varphi_i(x)\varphi_i(\xi)\,d\psi(x)$$

$$= \frac{a_{n-1}^*}{a_{n-1}\varphi_{n-1}(\xi)\,\varphi_{n-1}^*(\xi)} = \frac{1}{K_{n-1}(\xi;d\psi)}.$$

Similarly, if b is finite and we fix one point $\eta \geqq b$:

$$H_\eta = \frac{1}{K_{n-1}(\eta;d\psi)}.$$

The general case of m fixed points can be dealt with now very readily. With the notations (39), we get, letting $d\psi_1(x) = M(x)\,N(x)d\psi(x)$, with $M(x) \equiv \prod_{i=1}^{k}(x-\xi_i)$, $N(x) \equiv \prod_{i=k+1}^{m}(\eta_i - x)$:

$$H_{i,n} = \int_a^b \frac{M(x)\,N(x)\,\varphi_{n-m}(x;d\psi_1(x))\,d\psi(x)}{(x-x_{i,n})\,[M(x)\,N(x)\,\varphi_{n-m}(x;d\psi_1(x))]'_{x=x_{i,n}}} \qquad (i=1,2,\cdots,n).$$

For $1 \leqq i \leqq k$; $H_{i,n} \equiv H_{\xi i} = \int_a^b \frac{\frac{M(x)N(x)}{(x-\xi_i)}\varphi_{n-m}(x;d\psi_1(x))d\psi(x)}{M'(\xi_i)N(\xi_i)\varphi_{n-m}(\xi_i;d\psi_1(x))}$

$$= \frac{1}{M'(\xi_i)N(\xi_i)} \int_a^b \frac{\varphi_{n-m}(x;d\psi_1(x))d\psi_{1,i}(x)}{\varphi_{n-m}(\xi_i;d\psi_1(x))} = \frac{1}{M'(\xi_i)N(\xi_i)K_{n-m}(\xi_i;d\psi_{1,i}(x))}$$

$$\left(d\psi_{1,i}(x) = \frac{M(x)N(x)d\psi_1(x)}{x-\xi_i}\right).$$

For $k+1 \leqq i \leqq m$; $\quad H_{i,n} \equiv H_{\eta i}$

$$= \int_a^b \frac{\frac{M(x)N(x)}{\eta_i - x}\varphi_{n-m}(x;\,d\psi_1(x))d\psi(x)}{M(\eta_i)N'(\eta_i)\varphi_{n-m}(\eta_i;\,d\psi_1(x))} = \frac{1}{M(\eta_i)N'(\eta_i)} \int_a^b \frac{\varphi_{n-m}(x;d\psi_1(x))d\psi_{1,i}(x)}{\varphi_{n-m}(\eta_i;\,d\psi_1(x))}$$

$$= \frac{1}{M(\eta_i)N'(\eta_i)K_{n-m}(\eta_i;\,d\psi_{1,i}(x))} \left(d\psi_{1,i}(x) = \frac{M(x)N(x)d\psi_1(x)}{(\eta_i - x)}\right).$$

If, for example, (a, b) is finite, $m = 2$, $\xi_1 = a$, $\eta_1 = b$, then

$$H_a = \frac{1}{(b-a)K_{n-2}(a;(b-x)d\psi(x))}, \qquad H_b = \frac{1}{(b-a)K_{n-2}(b;(x-a)d\psi(x)}.$$

4. **The distribution of the zeros of Languerre and Hermite Polynomials.** Following a geometrical method the underlying idea of which is due to E. R. Neumann (for Laguerre Polynomials with $\alpha = 1$) [6], we shall find bounds for the zeros $\{x_{i,n}\}$ of Laguerre polynomials for any α (> 0) and incidentally for Hermite polynomials. Hereafter,

$$L_n(x; 0, \infty; e^{-x}x^{\alpha-1}) \equiv L_n(x) \equiv (-1)^n \varphi_n(x; 0, \infty; e^{-x}x^{\alpha-1}) \tag{40}$$

$$\equiv \frac{(-1)^n}{\sqrt{\Gamma(n+1)\Gamma(n+\alpha)}} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \binom{n+\alpha-1}{i} i! x^{n-i}.$$

The following relations, *which hold for any* $\alpha > 0$, are of fundamental importance and may be readily verified using the orthogonality properties of $L_n(x)$:

$$L'_{n+1}(x) - \sqrt{\frac{n+1}{n+\alpha}} L'_n(x) = -\sqrt{\frac{n+1}{n+\alpha}} L_n(x),$$

$$xL'_{n+1}(x) = (n+1) L_{n+1}(x) - \sqrt{(n+1)(n+\alpha)}\, L_n(x). \tag{41}$$

$$(\alpha > 0)$$

We have further, by (40),

$$\frac{L_n(x)}{L_n(0)} = 1 - \frac{n\Gamma(\alpha)}{\Gamma(\alpha+1)} x + \frac{n(n-1)\Gamma(\alpha)}{2!\,\Gamma(\alpha+2)} x^2 \pm \cdots + (-1)^n \frac{\Gamma(\alpha)}{\Gamma(\alpha+n)} x^n.$$

Hence

$$\frac{L_n(x)}{L_n(0)} > 0 \text{ for } x < \alpha/n,$$

$$S_n = \sum_{i=1}^{n} x_{i,n} = n(n+\alpha-1), \tag{42}$$

$$\alpha/n < x_{i,n} < n(n+\alpha-1) \qquad (i = 1, 2, \cdots, n). \tag{43}$$

In what follows $\mu_{i,n}$ and $\mu_{i,n+1}$ represent the abscissas of the extremal points—more briefly, "extremal abscissas"—of the curves $y = L_n(x)$ and $y = L_{n+1}(x)$ respectively. We derive from the differential equation (8, 9) satisfied by Laguerre polynomials:

$$\Lambda'_n(x) = -n^{\alpha-1} e^{-x} L_n(x) \qquad (\Lambda_n(x) \equiv x^\alpha e^{-x} L'_n(x)). \tag{44}$$

Hence the zeros of $\Lambda_n(x)$ are the extremal abscissas for $y = L_n(x)$ and conversely.

The following relations may be verified, making use of (8, 9, 40, 41):

$$x\Lambda''_n(x) + (x - \alpha + 1)\Lambda'_n(x) + n\Lambda_n(x) = 0, \tag{45}$$

$$x\sqrt{\frac{n+1}{n+\alpha}}\,\Lambda'_n(x) = n\left(\Lambda_{n+1}(x) - \sqrt{\frac{n+1}{n+\alpha}}\,\Lambda_n(x)\right). \tag{46}$$

$$\sqrt{(n+1)(n+\alpha)}\,\Lambda'_n(x) - n\Lambda'_{n+1}(x) = n\Lambda_{n+1}(x). \tag{47}$$

$$(48)\quad n\sqrt{\frac{n+\alpha}{n+1}}\,\Lambda_{n+1}(x)-(2n+\alpha-1-x)\,\Lambda_n(x)$$
$$+\sqrt{n(n+\alpha-1)}\,\Lambda_{n-1}(x)=0.$$

From (44, 46) it follows that the curve $y=\Lambda_{n+1}(x)$ intersects $y=\Lambda_n(x)$ at $x=x_{i,\,n}$ $(i=1,\,2,\,\cdots,\,n)$ (extremal abscissas for the latter curve) if $\alpha=1$, and to the right or left of $x_{i,\,n}$ according as $\alpha<1$ or $\alpha>1$. Formulae (44, 46, 47) give further:

$$n\int_{x_{i,\,n+1}}^{x_{i,\,n}}\Lambda_{n+1}(x)\,dx=\sqrt{(n+1)(n+\alpha)}\,\Lambda_n(x)-n\,\Lambda_{n+1}(x)\Big|_{x_{i,\,n+1}}^{x_{i,\,n}}$$
$$=\alpha\Lambda_{n+1}(x_{i,\,n})+(x_{i,\,n+1}-\alpha)\,\Lambda_{n+1}(x_{i,\,n+1}).$$

Let us take i even.[5] Then,

$$\Lambda_n(x_{i,\,n})\text{ is maximum},\qquad \Lambda_{n+1}(x_{i,\,n+1})\text{ is maximum};$$

$$\Lambda_n(x_{i,\,n})>\Lambda_n(x_{i,\,n+1})=\frac{n+\alpha-x_{i,\,n+1}}{\sqrt{(n+1)(n+\alpha)}}\,\Lambda_{n+1}(x_{i,\,n+1})\text{(by 44)},$$

$$n\int_{x_{i,\,n+1}}^{x_{i,\,n}}\Lambda_{n+1}(x)\,dx>\left(\frac{-\alpha x_{i,\,n+1}}{n+\alpha}+x_{i,\,n+1}\right)\Lambda_{n+1}(x_{i,\,n+1}),$$

$$\Lambda_{n+1}(x_{i,\,n+1})(x_{i,\,n}-x_{i,\,n+1})>\int_{n_{i,\,n+1}}^{x_{i,\,n}}\Lambda_{n+1}(x)\,dx,$$

$$n(x_{i,\,n}-x_{i,\,n+1})>\frac{-\alpha x_{i,\,n+1}}{n+\alpha}+x_{i,\,n+1},$$

$$(49)\qquad x_{i,\,n+1}<\frac{n+\alpha}{n+\alpha+1}\,x_{i,\,n},$$

$$(50)\qquad x_{i,\,n}<\frac{\alpha+i}{n+\alpha}\,x_{i,\,i}<\frac{\alpha+i}{n+\alpha}\,i(i+\alpha-1)\qquad(\alpha>0)\text{(see (42))}.$$

In particular, for $i=1$, and in conjunction with (43):

$$(51)\qquad \alpha/n<x_{1,\,n}<\alpha/n\,(1+\alpha).$$

We have further, by (42, 49):

$$n(n+\alpha-1)=\sum_{i=1}^{n}x_{i,\,n}<\frac{n+\alpha-1}{n+\alpha}\sum_{i=1}^{n-1}x_{i,\,n-1}+x_{n,\,n}$$
$$=x_{n,\,n}+\frac{n+\alpha-1}{n+\alpha}\,(n-1)(n+\alpha-2),$$

$$(52)\qquad x_{n,\,n}>3n+\alpha-5+\frac{2(\alpha+1)}{n+\alpha}\qquad(\alpha>0).$$

[5] If i is odd, we simply change the sign of $L_n(x)$.

It is known that $x_{n,n} < 4n + 2\alpha$; hence

$$3n + \alpha - 5 < x_{n,n} < 4n + 2\alpha \qquad (\alpha > 0).^{6} \tag{53}$$

In what follows, the figures, which for the sake of brevity have been omitted, are quite similar to those used by Neumann. From (8, 9, 40, 41) it is seen that the abscissa of each point of intersection of $y = L_n(x)$ and $y = L_{n+1}(x)$ is to the right or to the left of $x = \mu_{i,n+1}$ $(i = 1, 2, \cdots, n)$ according as $\alpha > 1$ or $\alpha < 1$. If $\alpha = 1$ we have Neumann's case: viz. the abscissas in question coincide with the points $x = \mu_{i,n+1}$.

Considerations similar to those employed by Neumann lead to the inequalities:

$$x_{1,n+1} < x_{1,n} < \mu_{1,n+1} < \begin{Bmatrix} x_{2,n+1} \\ \mu_{1,n} \end{Bmatrix} < x_{2,n} < \mu_{2,n+1} < \begin{Bmatrix} x_{3,n+1} \\ \mu_{2,n} \end{Bmatrix} < \cdots,$$

where $\mu_{i,n} \lesseqgtr x_{i+1,n+1}$ according as $\mu_{i,n} \lesseqgtr n + \alpha$. We also see that in the particular interval $(x_{i,n}, x_{i+1,n})$ containing α there will be no point of inflection of the curve $y = L_n(x)$. In every other interval $(x_{i,n}, x_{i+1,n})$ there will be just one such point of inflection situated to the right or left of $x = \mu_{i,n+1}$ according as $(x_{i,n}, x_{i+1,n})$ is to the left or right of $x = \alpha$.

Case 1. a) $x_{i+1,n} < \alpha$, $\alpha \leqq 1$. Here the above point of inflection lies between $\mu_{i,n+1}$ and $x_{i+1,n}$. Compare the area

$$\frac{(\mu_{i,n+1} - x_{i,n})^2}{2} L_n'(x_{i,n}) = \frac{(\mu_{i,n+1} - x_{i,n})^2}{2x_{i,n}} \sqrt{(n+1)(n+\alpha)}\, L_{n+1}(x_{i,n}) \quad \text{(by (41))}$$

of the triangle formed by the x-axis, the tangent to $y = L_n(x)$ at $x = x_{i,n}$ and the perpendicular to the x-axis at $x = \mu_{i,n+1}$ with the area under the curve $y = L_n(x)$ between $x_{i,n}$ and $\mu_{i,n+1}$:

$$\int_{x_{i,n}}^{\mu_{i,n+1}} L_n(x)\, dx = L_n(\mu_{i,n+1}) \left(\frac{1-\alpha}{n+1}\right) + \sqrt{\frac{n+\alpha}{n+1}}\, L_{n+1}(x_{i,n}) > \sqrt{\frac{n+\alpha}{n+1}}\, L_{n+1}(x_{i,n}) \qquad \text{(by (41))}.$$

Thus we get:

$$\frac{(\mu_{i,n+1} - x_{i,n})^2 \sqrt{(n+1)(n+\alpha)}}{2x_{i,n}} > \sqrt{\frac{n+\alpha}{n+1}}\,; \quad \mu_{i,n+1} - x_{i,n} > \sqrt{\frac{2x_{i,n}}{n+1}}$$

$$x_{i+1,n} - x_{i,n} > \sqrt{\frac{2x_{i,n}}{n+1}} \qquad (i = 1, 2, \cdots, n-1).$$

b) $x_{i,n} > \alpha$, $\alpha \leqq 1$. Here the above mentioned point of inflection lies in $(x_{i,n}, \mu_{i,n+1})$. The triangle considered is formed by the x-axis, the tangent to $y = L_n(x)$ at $x = x_{i+1,n}$ and the perpendicular to the x-axis at $x = \mu_{i,n}$. Its area

$$-\frac{(x_{i+1,n} - \mu_{i,n+1})^2 \sqrt{(n+1)(n+\alpha)}}{2x_{i+1,n}} L_{n+1}(x_{i+1,n}) \tag{54}$$

[6] For $\alpha = \frac{1}{2}, \frac{3}{2}$ see (63).

is greater than the corresponding area under the curve $y = L_n(x)$:

$$(55) \qquad \int_{\mu_{i,\,n+1}}^{x_{i+1,\,n}} L_n(x)\,dx = \left(\frac{\alpha-1}{n+1}\right) L_n(\mu_{i,\,n+1}) - \sqrt{\frac{n+\alpha}{n+1}}\, L_{n+1}(x_{i+1,\,n}).$$

We now seek to find a relation between $L_n(\mu_{i,\,n+1})$ and $L_{n+1}(x_{i+1,\,n})$. From the location of the above-mentioned point of inflection it is seen that

$$\frac{L_n(\mu_{i,\,n+1})}{x_{i+1,\,n} - \mu_{i,\,n+1}} < -L_n'(x_{i+1,\,n}) = -\frac{\sqrt{(n+1)(n+\alpha)}}{x_{i+1,\,n}} L_{n+1}(x_{i+1,\,n}) \qquad \text{(by (41))},$$

which, combined with (54, 55), leads to

$$-\frac{(x_{i+1,\,n} - \mu_{i,\,n+1})^2 \sqrt{(n+1)(n+\alpha)}}{2x_{i+1,\,n}} L_{n+1}(x_{i+1,\,n}) > \int_{\mu_{i,\,n+1}}^{x_{i+1,\,n}} L_n(x)\,dx$$

$$> -\left\{\frac{\alpha-1}{n+1}\,\frac{(x_{i+1,\,n} - \mu_{i,\,n+1})\sqrt{(n+1)(n+\alpha)}}{x_{i+1,\,n}} + \sqrt{\frac{n+\alpha}{n+1}}\right\} L_{n+1}(x_{i+1,\,n}).$$

Since $L_{n+1}(x_{i+1,\,n}) < 0$ for i even, we get further:

$$(x_{i+1,\,n} - \mu_{i,\,n+1})^2 - \frac{2(\alpha-1)(x_{i+1,\,n} - \mu_{i,\,n+1})}{n+1} > \frac{2x_{i+1,\,n}}{n+1} > \frac{2x_{i,\,n}}{n+1},$$

$$(56) \qquad x_{i+1,\,n} - \mu_{i,\,n+1} > \sqrt{\frac{2x_{i,\,n}}{n+1}} - \frac{1-\alpha}{n+1}.$$

Moreover, $L_n(\mu_{i,\,n+1})$ being the greatest value of $L_n(x)$ in $(x_{i,\,n},\ \mu_{i,\,n+1})$, we get (by (41)):

$$L_n(\mu_{i,\,n+1})(\mu_{i,\,n+1} - x_{i,\,n}) > \int_{x_{i,\,n}}^{\mu_{i,\,n+1}} L_n(x)\,dx > L_n(\mu_{i,\,n+1})\left(\frac{1-\alpha}{n+1}\right),$$

$$\mu_{i,\,n+1} - x_{i,\,n} > \frac{1-\alpha}{n+1},$$

and this, added to (56), gives:

$$(57) \qquad x_{i+1,\,n} - x_{i,\,n} > \sqrt{\frac{2x_{i,\,n}}{n+1}}.$$

Following Neumann, we now let

$$\rho_i = (n+1)\,x_{i,\,n}, \qquad \sigma_i = \frac{(i + 2\sqrt{\alpha} - 1)^2}{4}.$$

Then

$$\rho_i > \sigma_i \qquad \left(\alpha > \omega = \frac{(1+\sqrt{2})^2}{16} = .364\cdots\right) \qquad \text{(by induction)},$$

$$x_{i,\,n} > \frac{(i + 2\sqrt{\alpha} - 1)^2}{4(n+1)} \qquad (1 \geqq \alpha \geqq \omega;\ i = 1, 2, \cdots, n).$$

To find bounds for $x_{i,\,n}$ for $\alpha \leqq \omega$, let

$$\rho_i' = \frac{(n+1)\omega}{\alpha}\,x_{i,\,n}, \qquad \sigma_i' = \frac{(i + 2\sqrt{\omega} - 1)^2}{4}.$$

Then, as above, $\rho_i' > \sigma_i'$, and

$$x_{i,n} > \frac{(i + 2\sqrt{\omega} - 1)^2}{4(n+1)} \frac{\alpha}{\omega} \qquad (0 < \alpha \leqq \omega;\ i = 1, 2, \cdots, n).$$

Case 2. a) $x_{i+1,n} < \alpha$, $\alpha \geqq 1$. Here the above point of inflection lies in $(\mu_{i,n+1}, x_{i+1,n})$. In fact, it can be shown to be in $(\mu_{i,n}, x_{i+1,n})$. The area of the triangle formed by the x-axis, the tangent to $y = L_n(x)$ at $x = x_{i,n}$ and the perpendicular to the x-axis at $x = x_{i+1,n+1}$ is greater than $\int_{x_{i,n}}^{x_{i+1,n+1}} L_n(x)\,dx$. i.e., by (41),

$$\frac{(x_{i+1,n+1} - x_{i,n})^2 \sqrt{(n+1)(n+\alpha)}}{2x_{i,n}} L_{n+1}(x_{i,n}) > L_n(x_{i+1,n+1}) + \sqrt{\frac{n+\alpha}{n+1}}\, L_{n+1}(x_{i,n}),$$

$$\frac{(x_{i+1,n+1} - x_{i,n})^2 \sqrt{(n+1)(n+\alpha)}}{2x_{i,n}} > \sqrt{\frac{n+\alpha}{n+1}} \qquad (\text{since } L_n(x_{i+1,n+1}) > 0),$$

$$x_{i+1,n} - x_{i,n} > \sqrt{\frac{2x_{i,n}}{n+1}} \qquad (x_{i+1,n} < \alpha,\quad \alpha \geqq 1,\quad i = 1, 2, \cdots, n). \tag{58}$$

b) $x_{i,n} > \alpha$, $\alpha \geqq 1$. The point of inflection mentioned above lies in $(x_{i,n}, \mu_{i,n+1})$. The area of the triangle considered is

$$-\frac{(x_{i+1,n} - \mu_{i,n+1})^2 \sqrt{(n+1)(n+\alpha)}}{2x_{i+1,n}} L_{n+1}(x_{i+1,n})$$

so that here,

$$-\frac{(x_{i+1,n} - \mu_{i,n+1})^2 \sqrt{(n+1)(n+\alpha)}}{2x_{i+1,n}} L_{n+1}(x_{i+1,n}) > \int_{\mu_{i,n+1}}^{x_{i+1,n}} L_n(x)\,dx$$

$$= \left(\frac{\alpha - 1}{n+1}\right) L_n(\mu_{i,n+1}) - \sqrt{\frac{n+\alpha}{n+1}}\, L_{n+1}(x_{i+1,n})$$

$$\frac{(x_{i+1,n} - \mu_{i,n+1})^2 \sqrt{(n+1)(n+\alpha)}}{2x_{i+1,n}} > \sqrt{\frac{n+\alpha}{n+1}} \qquad (\text{by (41)}).$$

Thus we see that *formulae* (57, 58) hold for all $\alpha > 0$. We summarize our results as follows:

$$\frac{(i + 2\sqrt{\alpha} - 1)^2}{4(n+1)} < x_{i,n} < \frac{4(i+\alpha)^2}{n+\alpha} \qquad (\alpha \leqq \omega),\quad (i = 1, 2, \cdots, n) \tag{59}$$

$$\frac{(i + 2\sqrt{\omega} - 1)^2}{4(n+1)} \frac{\alpha}{\omega} < x_{i,n} < \frac{4(i+\alpha)^2}{n+\alpha} \qquad (\alpha \leqq \omega). \tag{60}$$

We now make use of (7), in order to obtain similar results for Hermite polynomials. Let $x_{i,2n+\sigma} (i = 1, 2, \cdots, 2n + \sigma;\ \sigma = 0, 1)$ represent the zeros of

Hermite polynomials of degree $2n+\sigma$, $x'_{i,n}$, $x''_{i,n}$ $(i=1,2,\cdots,n)$—those of the Laguerre polynomials of degree n, with $\alpha=\frac{1}{2},\frac{3}{2}$ respectively. Then,

$$(61)\qquad \begin{aligned} -x_{2n-1+i,\,2n} &= x_{i,\,2n} = \sqrt{x'_{i-n,\,n}} && (i=n+1,n+2,\cdots,2n),\\ -x_{2n+2-i,\,2n+1} &= x_{i,\,2n+1} = \sqrt{x''_{i-n-1,\,2n}} && (i=n+2,\cdots,2n+1); \end{aligned}$$

$$(62)^{7}\qquad \frac{\sqrt{2}}{2\sqrt{n}} < x_{n+1,\,2n} < \frac{\sqrt{3}}{2\sqrt{n}}, \qquad \frac{\sqrt{6}}{2\sqrt{n}} < x_{n+2,\,2n+1} < \frac{\sqrt{15}}{2\sqrt{n}} \qquad \text{(by (5))}.$$

Making use of (53, 59, 60, 61), we get:

$$(63)^{8}\qquad \sqrt{3n-\tfrac{9}{2}} < x_{2n,\,2n} < \sqrt{4n+1}, \qquad \sqrt{3n-\tfrac{7}{2}} < x_{2n+1,\,2n+1} < \sqrt{4n+3},$$

$$(64)\qquad \begin{aligned} \frac{i-n+\sqrt{2}-1}{2\sqrt{n+1}} < -x_{2n+1-i,\,2n} = x_{i,\,2n} < \frac{2\,(i-n+\frac{1}{2})}{\sqrt{n+\frac{1}{2}}} \\ (i=n+1,\cdots,2n), \\ \frac{i-n+\sqrt{6}-1}{2\sqrt{n+1}} < -x_{2n+2-i,\,2n+1} = x_{i,\,2n+1} < \frac{2\,(i-n+\frac{3}{2})}{\sqrt{n+\frac{3}{2}}} \\ (i=n+2,\cdots,2n+1). \end{aligned}$$

It is interesting to note that results quite similar to those given in (64) may be derived from Neumann's results without first considering the generalized Laguerre polynomials, but by making use of Markoff's theorem [14]. In fact, this theorem yields the inequality

$$\frac{dx_{i,\,n}}{d\alpha} > 0 \qquad (i=1,2,\cdots,n;\ \alpha>0),$$

which enables us to compare $x'_{i,\,n}$ $(\alpha=\frac{1}{2})$, $x''_{i,\,n}$ $(\alpha=\frac{3}{2})$ with $x_{i,\,n}$ $(\alpha=1)$. Thus we get for Hermite polynomials:

$$\begin{aligned} -x_{2n+1-i,\,2n} = x_{i,\,2n} = \sqrt{x'_{i-n,\,n}} < \sqrt{x_{i-n,\,n}} < \frac{2\,(i-n+1)}{\sqrt{n+1}} \\ (i=n+1,\cdots,2n), \\ -x_{2n+2-i,\,2n+1} = x_{i,\,2n+1} = \sqrt{x''_{i-n-1,\,n}} > \sqrt{x_{i-n-1,\,n}} > \frac{i-n}{2\sqrt{n+1}} \\ (i=n+2,\cdots,2n+1); \end{aligned}$$

and since $x_{i,\,2n} > x_{i,\,2n+1}$ $(1 \leqq i \leqq 2n)$,

$$\frac{i-n}{2\sqrt{n+1}} < -x_{2n+1-i,\,2n} = x_{i,\,2n} < \frac{2(i-n+1)}{\sqrt{n+1}} \qquad (i=n+2,\cdots,2n),$$

[7] A. Wiman [12], employing a different method, finds similar results.

[8] The bounds for $x_{n,\,n}$ have been much improved recently by S. C. Van Veen [13] on the basis of a less elementary analysis.

(65)

$$\frac{i-n}{2\sqrt{n+1}} < -x_{2n+2-i,\,2n+1} = x_{i,\,2n+1} < \frac{2(i-n+1)}{\sqrt{n+1}} \quad (i = n+2, \cdots, 2n+1).$$

Bibliography

1. Gauss. Methodus nova integralum valores per approximationem inveniendi, Werke, v. III, (165–96).
2. Stieltjes. Quelques recherches sur la théorie des quadratures dites mécaniques, Oeuvres, v. 1, (377–94).
3. J. V. Uspensky. a) On the Convergence of Quadrature Formulas Related to an Infinite Interval, Trans. Am. Math. Soc., v. 30, 1928 (542–59).
   b) On the Convergence of Quadrature Formulas Between Infinite Limits, Bull. Russ. Acad. Science, 1916 (in Russian).
4. J. Chokhate (J. Shohat). Sur la convergence des quadratures mécaniques dans un intervalle infini, Comptes Rendus, v. 186, 1928 (344–6).
5. N. Sonine. On the Determination of Limiting Values of Definite Integrals, Memoirs Russ. Acad. Sciences, v. 69 (1892) (in Russian).
6. E. R. Neumann. Beiträge zur Kenntnis der Laguerreschen Polynome, Jahresbericht der Deutchen math. Vereinigung, v. 30, 1921 (15–35).
7. J. Shohat. On a Certain Formula of Mechanical Quadratures with Non-Equidistant Ordinates, Trans. Am. Math. Soc., v. 31, 1929 (448–463).
8. M. Plancherel et W. Rotach. Sur les valeurs asymptotiques des polynomes d'Hermite, Commentarii Mathematici Helvetii, v. 1, 1929 (227–54).
9. E. Kogbetliantz. Recherches sur la sommabilité des séries d'Hermite, Annales l'École Normale Supérieure, v. 49 (137–221).
10. Darboux. Mémoire sur l'approximation des fonctions des grands nombres, Journ. des Math., III sér., t. IV, 1878 (5–57, 377–416).
11. J. Shohat. On the Asymptotic Properties of a Certain Class of Tchebycheff Polynomials, Proc. Int. Math. Cong., Toronto, v. 1, (611–618).
12. A. Wiman. Ableitungen ganzer Funktionen, Math. Annalen, v. 104 (169–81).
13. S. C. Van Veen. Asymptotische Entwicklung der Hermiteschen Funktionen, Math. Annalen, v. 105, 1931 (408–36).
14. A. Markoff. Math. Annalen, t. 27, 1886 (177–182).

The University of Pennsylvania.

Added in Proof (July 9, 1934):

Since the time this paper was sent in for publication there has appeared a paper [*Wolfgang Hahn*—Die Nullstellen der Laguerreschen und Hermiteschen Polynome.—Schriften des Math. Sem. und des Inst. für angewandte Math. der Univ. Berlin, Band 1.] in which results similar to those gotten here for the zeros of Hermite and Laguerre Polynomials are derived by similar methods. No applications to Mechanical Quadratures are made however [cf. also *J. Shohat and C. Winston*; On Mechanical Quadratures, Rend. Del Circ. Mat. Di Palermo, Tomo LVIII, Anno 1934–XII.]

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

SUR LES NOMBRES DE BETTI LOCAUX[1]

PAR EDUARD ČECH

(Received November 23, 1933)

Soit R un espace topologique arbitraire; soit a un point donné de R; soit $k = 0, 1, 2, \cdots$. Au Chap. I, je définis le $k^{\text{ième}}$ *nombre de Betti local de R au point a*, désigné par $\beta_k(a, R)$. La signification du nombre $\beta_0(a, R)$ est très simple (v. n°3). Si $k \geqq 1$ et si l'espace R est un polyèdre, le nombre $\beta_k(a, R)$ coïncide, comme on pourrait le démontrer sans peine, avec le nombre maximum des $(k-1)$-cycles en a lin. indépendants, au sens de *M. E. R. van Kampen*.[2]

Au Chap. II, je prouve deux théorèmes d'addition pour les nombres de Betti locaux. Ces théorèmes pourraient être sans peine généralisés.

Au Chap. III, je donne une nouvelle forme aux axiomes de ma théorie des variétés, exposée récemment dans ce Journal.[3] Dans la nouvelle forme, ces axiomes reposent sur la notion des nombres de Betti locaux et sur celle de la connexité locale d'ordre supérieur que j'étudie dans un autre Mémoire (cité au n°4). En se basant sur les nouveaux axiomes, on pourrait un peu simplifier quelques démonstrations dans *Variétés*.

Soit R un espace topologique arbitraire; soit S un sous-ensemble fermé de R; soit a un point donné de S; soit $k = 0, 1, 2, 3, \cdots$. Au Chap. IV, je définis ce qu'on pourrait appeler le $k^{\text{ième}}$ *nombre de Betti extérieur de S au point a*, désigné par $\alpha_k(a, R - S)$. Si l'espace R est régulier et localement connexe, l'ensemble S coupe l'espace R au point a localement en $\alpha_0(R - S) + 1$ régions (v. n°25); autrement dit S détermine en a une "lokale Zerschneidung" de R d'ordre $\alpha_0(R - S) + 1$ au sens de M. K. Zarankiewicz.[4]

Au Chap. V, je donne une localisation du théorème de dualité: *Si S est un sous-ensemble fermé de l'espace euclidien R_n à n dimensions*,[5] *on a en chaque point*

---

[1] J'ai exposé quelques résultats de ce Mémoire dans une conférence que j'ai faite le 5 juillet 1933 dans le Math. Kolloquium de M. Menger.

Après avoir terminé ce travail j'ai pris connaissance de deux notes de M. P. Alexandroff: *Sur les propriétés locales des ensembles fermés* (C. R. Paris 198, p. 227, 15 janvier 1934) et *Les groupes de Betti en un point* (ibidem, p. 315, 22 janvier 1934) qui semblent avoir beaucoup de points de contact avec mes recherches.

(Note de la rédaction: Un mémoire étendu de M. P. Alexandroff sur ces questions paraîtra dans le numéro d'octobre des Annals of Mathematics.)

[2] *Die kombinatorische Topologie und die Dualitätssätze*, den Haag 1929, p. 26.

[3] *Théorie générale des variétés et de leurs théorèmes de dualité*, Annals of Math., t. 34, 1933, p. 621–730. Cité: *Variétés*.

[4] *Über die lokale Zerschneidung der Ebene*, Monatshefte f. Math. u. Phys., t. 39, 1932, p. 371.

[5] Le théorème est même démontré pour des espaces plus généraux que R_n.

a de S: $\beta_r(a, S) = \alpha_q(a, R - S)$, $0 \leqq p \leqq n-1$, $q = n - p - 1$. En particulier l'ordre de coupure locale est un invariant topologique (local) de S.

Soit S un sous-ensemble fermé de R_{n-1}; soit a un point de S. Si a est un point intérieur de S, on a $\beta_{n-1}(a, S) = 1$; si a est un point frontière de S, on a $\beta_{n-1}(a, S) = 0$ (Cf. n°7, 6° et 7°). Donc, *si l'on immerge S dans un R_n, a est un point de coupure locale si et seulement si c'est un point intérieur de S, et l'ordre de coupure locale est alors égal à deux.*

Un cas particulier ($R = R_n$, $k = n - 1$) du théorème du n°5 est: *Soient A et B deux sous-ensembles fermés de R_n; soit a un point de AB; Supposons que a soit un point intérieur de $A + B$, mais un point frontière de A et de B. Alors AB coupe R_n localement en a.*

Un cas particulier ($R = R_n$, $k = n - 1$) du théorème du n°6 est: *Soient A et B deux sous-ensembles fermés de R_n; soit a un point de AB. Si AB coupe R_n localement en a, tandis que ni A ni B ne coupe R_n localement en a, alors a est un point intérieur de $A + B$.*

Un cas particulier ($R = R_2$, $k = 0$) du théorème du n°5 est: *Soient A et B deux sous-ensembles fermés du plan; soit a un point de AB. Si ni A ni B ne coupent localement le plan en a, tandis que $A + B$ coupe localement le plan en a, alors ou bien a est un point isolé de AB (et alors $A + B$ coupe le plan localement en a en deux régions), ou bien chaque entourage de a contient une infinité de composantes de AB.*

Un cas particulier ($R = R_2$, $k = 0$) du théorème du n°5 est: *Soient A et B deux sous-continus du plan; soit a un point de AB. Si a est un point isolé de AB, alors $A + B$ coupe le plan localement en a. Si chaque entourage de a contient une infinité de composantes de AB, alors a est un point de coupure locale du plan d'ordre infini pour $A + B$.*

Les deux derniers théorèmes constituent dans un certain sens une localisation des deux théorèmes classiques de *Janiszewski*.[6] Une localisation entièrement différente de ces théorèmes de *Janiszewski* se trouve au Mémoire cité au n°4.

Je suppose dans tous ce Mémoire que les coefficients des cycles et des homologies appartiennent à $\mathfrak{R}$, où $\mathfrak{R}$ signifie ou bien l'ensemble de tous les nombres rationnels ou bien l'ensemble de tous les entiers réduits mod p, p étant un nombre premier donné d'avance.

Pour les domaines $\mathfrak{R}$ ici considérés, les théorèmes locaux de dualité constituent une généralisation du *allgemeiner dimensionstheoretischer Rechtfertigungssatz* de *M. P. Alexandroff*.[7]

Quant à la théorie de l'homologie, je m'appuie sur mon Mémoire *Théorie générale de l'homologie dans un espace quelconque*,[8] cité: *Homologie*. Comme dans *Variétés* (p. 622, I), j'écrit C^p(p. ex.) au lieu de $\{C^p(\mathfrak{U})\}$ (*Homologie*, II, 20). Donc C^p est l'ensemble de tous les $C^p(\mathfrak{U})$, $\mathfrak{U}$ parcourant les réseaux[9] dans R.

---

[6] *Sur les coupures locales faites par les continus*, Prace mat.-fiz., t. 29, 1913, pp. 11–63.
[7] *Dimensionstheorie*, Math. Annalen, t. 106, pp. 161–238 (v. p. 208).
[8] Fund. Math., t. 19, 1932, pp. 149–183.
[9] La famille fondamentale de réseaux (*Homologie*, II, 1) est dans tout ce Mémoire celle de tous les réseaux ouverts (*Homologie*, III, 2).

Soit $\mathfrak{U}$ un réseau; soit $C^\circ(\mathfrak{U}) = \sum_{i=1}^{\alpha_0} r_i U_i (r_i \in \mathfrak{R}, U_i \in \mathfrak{U})$ un $(0, \mathfrak{U})$-cycle. Je pose $J[C^\circ(\mathfrak{U})] = \sum_{i=1}^{\alpha_0} r_i$. Si C° est un $(0, R)$-cycle absolu, le nombre $J[C^\circ(\mathfrak{U})]$ est, comme on le voit sans peine, indépendant du réseau $\mathfrak{U}$; je désigne ce nombre par $J(C^\circ)$.

I.

1. Soit R un espace topologique (*Homologie*, III, 1). Soit S un sous-ensemble fermé de R. Soit a un point donné de S. Soit $k = 0, 1, 2, \cdots$. Les lettres U, V, W désignent des entourages[10] de a dans R.

Soit $U \supset V$. Soit $\mathfrak{B}_k(V, U; S)$ l'ensemble de tous les (k, R)-cycles mod $(S - U)$ dans S; deux éléments C^k et D^k de $\mathfrak{B}_k(V, U; S)$ seront considérés comme égaux si et seulement si $C^k \sim D^k \bmod (S - V)$ dans S. L'ensemble $\mathfrak{B}_k(V, U; S)$ est un module (*Homologie*, I, 1). Désignons par $\beta_k(V, U; S)$ le rang (*Homologie*, I, 3) de ce module si ce rang est fini; dans le cas contraire posons $\beta_k(V, U; S) = \infty$.

Pour $W \subset V \subset U$ on a évidemment $\beta_k(W, U; S) \leqq \beta_k(V, U; S)$.

Il en résulte que, l'entourage U de a étant donné, le nombre $\beta_k(V, U; S)$ a une valeur fixe (indépendante de V) pour tous les voisinages $V \subset U$ de a suffisamment petits; désignons par $\beta_k(a, U; S)$ cette valeur fixe.

Pour $W \subset V \subset U$ on a évidemment $\beta_k(W, V; S) \geqq \beta_k(W, U; S)$. On en déduit sans peine que $\beta_k(a, V; S) \geqq \beta_k(a, U; S)$ pour $V \subset U$. Par suite trois cas sont à distinguer:

1°. Il existe un nombre fini m $(= 0, 1, 2, \cdots)$ tel que $\beta_k(a, U; S) = m$ pour tous les entourages U de a suffisamment petits; dans ce cas on pose $\beta_k(a, S) = m$.

2°. Le nombre $\beta_k(a, U; S)$ est fini pour tous les entourages U de a, mais si m est un nombre fini arbitrairement donné, on a $\beta_k(a, U; S) > m$ pour tous les entourages U de a suffisamment petits; dans ce cas on pose $\beta_k(a, S) = \omega$.

3°. On a $\beta_k(a, U; S) = \infty$ pour tous les entourages U de a suffisamment petits; dans ce cas on pose $\beta_k(a, S) = \infty$.

2. De *Homologie*, III, 3–11 on déduit sans peine que le nombre $\beta_k(a, S)$ ne dépend nullement de l'espace R, mais seulement du nombre k, du point a et (localement) de l'espace S.

3. *Si l'espace R est régulier au point a,*[11] *le nombre $\beta_0(a, S)$ a toujours une des trois valeurs* 0, 1, ∞: *On a $\beta_0(a, S) = 0$ si et seulement s'il existe un entourage U de a tel que chaque composante de S a un point commun avec $R - U$. On a*

---

[10] *Entourage* d'un point ou d'un sous-ensemble est un ensemble ouvert contenant le point ou le sous-ensemble considéré.

[11] Cela signifie, comme on sait, que chaque entourage U de a contient un entourage V de a tel que $\bar{V} \subset U$.

$\beta_0(a, S) = 1$ si et seulement si a est un point isolé de l'espace S. On a $\beta_0(a, S) = \infty$ si et seulement si, pour chaque entourage U de a, une infinité de composantes de S sont contenues dans SU.

DÉMONSTRATION. Si a est un point isolé de S, on a évidemment $\beta_0(a, S) = 1$, car $\beta_0(V, U; S) = 1$ si U est tel que $SU = (a)$. S'il existe un entourage U de a tel que chaque composante de S rencontre $R - U$, d'après *Homologie*, III, 16 on a $\beta_0(W, V; S) = 0$ si $V \subset U$, d'où $\beta_0(a, S) = 0$. Ces deux cas étant exclus, pour chaque entourage U de a l'ensemble US contient une infinité de composantes de S et donc aussi une infinité de quasicomposantes de S. L'espace R étant régulier au point a, il existe un entourage V de a tel que $\bar{V} \subset U$. L'ensemble VS contient une infinité de quasicomposantes de S; chaque telle quasicomposante est évidemment essentielle mod $(S - U)$ au sens de *Homologie*, III, 15, d'où $\beta_0(V, U; S) = \infty$ d'après *Homologie*, III, 17. Donc $\beta_0(a, S) = \infty$.

II.

4. LEMME. *Soit R un espace complètement normal* (*Homologie*, III, 19). *Soient φ et χ deux sous-ensembles fermés de R. Soit $\mathfrak{U}$ un réseau donné dans R. Il existe un affinement $\mathfrak{V}$ de $\mathfrak{U}$ tel que, si une $(p, \mathfrak{V})$-chaîne est contenue dans φ et dans χ* (*Homologie*, II, 5), *elle est aussi contenue dans $\varphi\chi$.*

La démonstration se trouve au n°12 de mon Mémoire *Sur la connexité locale d'ordre supérieur* (à paraître dans Compositio Mathematica).

5. *Soit R un espace complètement normal. Soit $R = A + B$, A et B étant des sous-ensembles fermés de R. Soit $k = 0, 1, 2, \cdots$. Soit $a \in AB$. Soit $\beta_{k+1}(a, A) = \beta_{k+1}(a, B) = 0$. Alors $\beta_k(a, AB) \geqq \beta_{k+1}(a, R)$.*

DÉMONSTRATION. On voit sans peine qu'il suffit de prouver que, U étant un entourage donné de a, l'inégalité $\beta_{k+1}(a, U; R) \geqq m$ $(= 0, 1, 2, \cdots)$ entraîne que $\beta_k(a, U; AB) \geqq m$. Soit donc $\beta_{k+1}(a, U; R) \geqq m$ et soit $V \subset U$ un entourage de a. Evidemment, il suffit d'en déduire que $\beta_k(V, U; AB) \geqq m$.

Puisque $\beta_{k+1}(a, A) = 0$, on a $\beta_{k+1}(a, V; A) = 0$. Donc il existe un entourage $W_1 \subset V$ de a tel que $\beta_{k+1}(L_1, V; A) = 0$ pour chaque entourage $L_1 \subset W_1$ de a. Pareillement on voit qu'il existe un entourage $W_2 \subset V$ de a tel que $\beta_{k+1}(W_1', V; B) = 0$ pour chaque entourage $W_2' \subset W_2$ de a. Posons $W = W_1 W_2$ de manière que $\beta_{k+1}(W, V; A) = \beta_{k+1}(W, V; B) = 0$.

Puisque $\beta_{k+1}(a, U; R) \geqq m$, on a $\beta_{k+1}(W, U; R) \geqq m$. Donc il existe des $(k + 1, R)$-cycles C_i^{k+1} mod $(R - U)$ $(1 \leqq i \leqq m)$ tels que l'homologie $\sum_{i=1}^{m} r_i C_i^{k+1} \sim 0$ mod $(R - W)$ $(r_i \in \mathfrak{R})$ entraîne que $r_1 = \cdots = r_m = 0$.

Donc il existe un réseau $\mathfrak{U}_1$ tel que l'homologie $\sum_{i=1}^{m} r_i C_i^{k+1}(\mathfrak{U}_1) \sim 0$ mod $(R - W)$ entraîne que $r_1 = \cdots = r_m = 0$.

Soit $\mathfrak{U}_2$ un affinement de $\mathfrak{U}_1$ normal (*Homologie*, II, 15 et 16) rel. aux cycles mod $(B - V)$ dans B. Déterminons un affinement $\mathfrak{U}_3$ de $\mathfrak{U}_2$ d'après 4, en y posant $\varphi = B$, $\chi = R - V$. Soit $\mathfrak{U}_4$ un affinement de $\mathfrak{U}_3$ normal rel. aux cycles mod $(A - V)$ dans A. Déterminons un affinement $\mathfrak{U}_5$ de $\mathfrak{U}_4$ d'après 4, en y posant $\varphi = A$, $\chi = R - V$. Soit $\mathfrak{U}_6$ un affinement de $\mathfrak{U}_5$ normal rel. aux cycles mod $(AB - V)$ dans AB. Déterminons un affinement $\mathfrak{U}_7$ de $\mathfrak{U}_6$ d'après 4, en y posant $\varphi = AB$, $\chi = R - U$ ainsi qu'un affinement $\mathfrak{U}_8$ de $\mathfrak{U}_7$ de nouveau d'après 4, mais en y posant $\varphi = A$, $\chi = B$. Soit $\pi_{21} = \text{Pr.}\ (\mathfrak{U}_{21}\ \mathfrak{U}_1), \cdots,$ $\pi_{87} = \text{Pr.}\ (\mathfrak{U}_8, \mathfrak{U}_7)_1\ \pi_{31} = \pi_{21}\,\pi_{321} \cdots$.

Comme $A + B = R$, on peut poser $C_i^{k+1}(\mathfrak{U}_8) = C_{i1}^{k+1}(\mathfrak{U}_8) - C_{i2}^{k+1}(\mathfrak{U}_8)$ $(1 \leqq i \leqq m)$, où $C_{i1}^{k+1}(\mathfrak{U}_8) \subset A$, $C_{i2}^{k+1}(\mathfrak{U}_8) \subset B$. Soit $\Gamma_i^k(\mathfrak{U}_8)$ cette partie de la $(k, \mathfrak{U}_8)$-chaîne $FC_{i1}^{k+1}(\mathfrak{U}_8)$ dont les simplexes sont $\neq 0$ mod $(R - U)$. Comme $C_i^{k+1}(\mathfrak{U}_8) \to 0$ mod $(R - U)$, on a $FC_{i1}^{k+1}(\mathfrak{U}_8) = FC_{i2}^{k+1}(\mathfrak{U}_8)$ mod $(R - U)$; donc $\Gamma_i^k(\mathfrak{U}_8)$ est aussi cette partie de la chaîne $FC_{i2}^{k+1}(\mathfrak{U}_8)$ dont les simplexes sont $\neq 0$ mod $(R - U)$. Comme $C_{i1}^{k+1}(\mathfrak{U}_8) \subset A$, on a $\Gamma_i^k(\mathfrak{U}_8) \subset A$. Comme $C_{i2}^{k+1}(\mathfrak{U}_8) \subset B$, on a $\Gamma_i^k(\mathfrak{U}_8) \subset B$. D'après la définition de $\mathfrak{U}_8$ il en résulte que $\Gamma_i^k(\mathfrak{U}_8) \subset AB$, d'où $F\Gamma_i^k(\mathfrak{U}_8) \subset AB$. Comme $\Gamma_i^k(\mathfrak{U}_8) = FC_{i1}^{k+1}$ $(\mathfrak{U}_8)$ mod $(R - U)$, on a $F\Gamma_i^k(\mathfrak{U}_8) \subset R - U$. Donc $F\pi_{87}\Gamma_i^k(\mathfrak{U}_8) \subset AB$, $F\pi_{87}$ $\Gamma_i^k(\mathfrak{U}_8) \subset R - U$, d'où $F\pi_{87}\ \Gamma_i^k(\mathfrak{U}_8) \subset AB - U$ d'après la définition de $\mathfrak{U}_8$. Donc $\pi_{86}\ \Gamma_i^k(\mathfrak{U}_8)$ est un $(k, \mathfrak{U}_6)$-cycle mod $(AB - U)$ dans AB. D'après la définition de $\mathfrak{U}_6$, $\pi_{85}\ \Gamma_i^k(\mathfrak{U}_8)$ est un $(k, \mathfrak{U}_5)$-cycle essentiel mod $(AB - U)$ dans AB.

Il ne reste qu'à démontrer que les $(k, \mathfrak{U}_5)$-cycles $\pi_{85}\ \Gamma_i^k(\mathfrak{U}_8)$ $(1 \leqq i \leqq m)$ ne sont liés par aucune homologie mod $(AB - V)$ dans AB, car alors (*Homologie*, II, 28) $\beta_k(V, U; AB) \geqq m$, c. g. f. d.

Soit donc $\sum_{i=1}^{m} r_i\ \pi_{85}\ \Gamma_i^k(\mathfrak{U}_8) \sim 0$ mod $(AB - V)$ dans AB $(r_i \in \mathfrak{R})$; on doit prouver que $r_1 = \cdots = r_m = 0$. L'homologie qui vient d'être écrite signifie qu'il existe une $(k + 1, \mathfrak{U}_5)$-chaîne $D^{k+1}(\mathfrak{U}_5)$ dans AB telle que

(1) $FD^{k+1}(\mathfrak{U}_5) = \pi_{85} \sum_{i=1}^{m} r_i\ \Gamma_i^k(\mathfrak{U}_8)$ mod $(AB - V)$. Or $\sum_{i=1}^{m} r_i\ C_{i1}^{k+1}(\mathfrak{U}_8) \to$ $\sum_{i=1}^{m} r_i\ \Gamma_i^k(\mathfrak{U}_8)$ mod $(R - U)$. Donc la $(k + 1, \mathfrak{U}_5)$-chaîne $\pi_{85} \sum_{i=1}^{m} r_i\ C_{i1}^{k+1}(\mathfrak{U}_8) -$ $D^{k+1}(\mathfrak{U}_5)$ est située dans A et sa frontière est située dans $(R - V)$.

D'après la définition de $\mathfrak{U}_5$, on a donc $F\pi_{85} \sum_{i=1}^{m} r_i\ C_{i1}^{k+1}(\mathfrak{U}_8) - FD^{k+1}(\mathfrak{U}_5) \subset$ $A - V$. Donc $\pi_{84} \sum_{i=1}^{m} r_i\ C_{i1}^{k+1}(\mathfrak{U}_8) - \pi_{54}\ D^{k+1}(\mathfrak{U}_5)$ est un $(k + 1, \mathfrak{U}_4)$-cycle mod $(A - V)$ dans A. D'après la définition de $\mathfrak{U}_4$, $\pi_{83} \sum_{i=1}^{m} r_i\ C_{i1}^{k+1}(\mathfrak{U}_8) -$ $\pi_{53}\ D^{k+1}(\mathfrak{U}_5)$ est un $(k + 1, \mathfrak{U}_3)$-cycle essentiel mod $(A - V)$ dans A et par suite

(2) $\pi_{81} \sum_{i=1}^{m} r_i\, C_{i1}^{k+1}(\mathfrak{U}_8) - \pi_{51}\, D^{k+1}(\mathfrak{U}_5)$ est un $(k+1, U_1)$-cycle essentiel mod $(A - V)$ dans A. De plus $\sum_{i=1}^{m} r_i\, C_{i2}^{k+1}(\mathfrak{U}_8) \to \sum_{i=1}^{m} r_i\, \Gamma_i^k(\mathfrak{U}_8)$ mod $(R - U)$. On voit donc, en tenant compte aussi de (1), que la $(k+1, \mathfrak{U}_3)$-chaîne $\pi_{83} \sum_{i=1}^{m} r_i\, C_{i2}^{k+1}(\mathfrak{U}_8) - \pi_{53}\, D^{k+1}(\mathfrak{U}_5)$ est située dans B et que sa frontière est située dans $(R - V)$. D'après la définition de $\mathfrak{U}_3$, on a donc $F\pi_{83} \sum_{i=1}^{m} r_i\, C_{i2}^{k+1}(\mathfrak{U}_8) - F\pi_{53}\, D^{k+1}(\mathfrak{U}_5) \subset B - V$. Donc $\pi_{82} \sum_{i=1}^{m} r_i\, C_{i2}^{k+1}(\mathfrak{U}_8) - \pi_{52}\, D^{k+1}(\mathfrak{U}_5)$ est un $(k+1, \mathfrak{U}_2)$-cycle mod $(B - V)$ dans B. D'après la définition de $\mathfrak{U}_2$

(3) $\pi_{81} \sum_{i=1}^{m} r_i\, C_{i2}^{k+1}(\mathfrak{U}_8) - \pi_{51}\, D^{k+1}(\mathfrak{U}_5)$ est un $(k+1, \mathfrak{U}_1)$-cycle essentiel mod $(B - V)$ dans B.

Comme $\beta_{k-1}(W, V; A) = 0$, il résulte de (2) (en vertu de *Homologie*, II, 28) que $\pi_{81} \sum_{i=1}^{m} r_i C_{i1}^{k+1}(\mathfrak{U}_8) - \pi_{51} D^{k+1}(\mathfrak{U}_5) \sim 0$ mod $(A - W)$ dans A. Comme $\beta_{k+1}(W, V; B) = 0$, il résulte de (3) que $\pi_{81} \sum_{i=1}^{m} r_i C_{i2}^{k+1}(\mathfrak{U}_8) - \pi_{51} D^{k+1}(\mathfrak{U}_5) \sim 0$ mod $(B - W)$ dans B. Donc il existe des $(k+2, \mathfrak{U}_1)$-chaînes $E_1^{k+2}(\mathfrak{U}_1) \subset A$ et $E_2^{k+2}(\mathfrak{U}_1) \subset B$ telles que

$$E_1^{k+2}(\mathfrak{U}_1) \to \pi_{81} \sum_{i=1}^{m} r_i C_{i1}^{k+1}(\mathfrak{U}_8) - \pi_{51} D^{k+1}(\mathfrak{U}_5) \text{ mod } (A - W),$$

$$E_2^{k+2}(\mathfrak{U}_1) \to \pi_{81} \sum_{i=1}^{m} r_i C_{i2}^{k+1}(\mathfrak{U}_8) - \pi_{51} D^{k+1}(\mathfrak{U}_5) \text{ mod } (B - W).$$

Donc $E_1^{k+2}(\mathfrak{U}_1) - E_2^{k+2}(\mathfrak{U}_1) \to \pi_{81} \sum_{i=1}^{m} r_i C_i^{k+1}(\mathfrak{U}_8)$ mod $(R - W)$, d'où $\pi_{81} \sum_{i=1}^{m} r_i\, C_i^{k+1}(\mathfrak{U}_8) \sim 0$ mod $(R - W)$. Or, C_i^{k+1} étant un $(k+1, R)$-cycle mod $(R - U)$, on a $\pi_{81} C_i^{k+1}(\mathfrak{U}_8) \sim C_i^{k+1}(\mathfrak{U}_1)$ mod $(R - U) \subset (R - W)$. Donc $\sum_{i=1}^{m} r_i C_i^{k+1}(\mathfrak{U}_1) \sim 0$ mod $(R - W)$. D'après la définition de $\mathfrak{U}_1$, il en résulte que $r_1 = \cdots = r_m = 0$, c.q.f.d.

6. *Soit R un espace complètement normal. Soit $R = A + B$, A et B étant des sous-ensembles fermés de R. Soit $a \in AB$. Soit $k = 0, 1, 2, \cdots$. Soit $\beta_k(a, A) = \beta_k(a, B) = 0$. Alors $\beta_k(a, AB) \leqq \beta_{k+1}(a, R)$.*

DÉMONSTRATION. On voit sans peine qu'il suffit de prouver que, U étant un entourage donné de a, il existe un entourage $V \subset U$ de a tel que l'inégalité $\beta_k(a, U; AB) \geqq m$ $(= 0, 1, 2, \cdots)$ entraîne que $\beta_{k+1}(a, V; R) \geqq m$. Soit donc $\beta_k(a, U; AB) \geqq m$. Comme $\beta_k(a, A) = \beta_k(a, B) = 0$, on a

$$\beta_k(a, U; A) = \beta_k(a, U; B) = 0.$$

Donc on peut choisir l'entourage $V \subset U$ de a de manière que

$$\beta_k(V, U; A) = \beta_k(V, U; B) = 0$$

Il s'agit de prouver que, W étant un entourage de a contenu dans V, on a $\beta_{k+1}(W, V; R) \geqq m$.

Puisque $\beta_k(a, U; AB) \geqq m$, on a $\beta_k(W, U; AB) \geqq m$. Donc il existe des (k, R)-cycles Γ_i^k mod $(AB - U)$ dans AB $(1 \leqq i \leqq m)$ tels que l'homologie $\sum_{i=1}^{m} r_i \Gamma_i^k \sim 0 \bmod (AB - W)$ dans AB $(r_i \in \mathfrak{R})$ entraîne que $r_1 = \cdots = r_m = 0$.

Donc il existe un réseau $\mathfrak{U}_1$ tel que l'homologie $\sum_{i=1}^{m} r_i \Gamma_i^k(\mathfrak{U}_1) \sim 0 \bmod (AB - W)$ dans AB entraîne que $r_1 = \cdots = r_m = 0$. Déterminons d'après 4 d'abord un affinement $\mathfrak{U}_2$ de $\mathfrak{U}_1$ et ensuite un affinement $\mathfrak{U}_3$ de $\mathfrak{U}_2$, en y posant dans le premier cas $\varphi = AB$, $\chi = R - W$ et dans le second $\varphi = A$, $\chi = B$. Soit $\mathfrak{U}_4$ un affinement de $\mathfrak{U}_3$ normal rel. aux cycles mod $(R - V)$ dans R. Soit $\pi_{21} = \text{Pr.}(\mathfrak{U}_1, \mathfrak{U}_2)$, $\pi_{32} = \text{Pr.}(\mathfrak{U}_3, \mathfrak{U}_2)$, $\pi_{43} = \text{Pr.}(\mathfrak{U}_4, \mathfrak{U}_3)$, $\pi_{42} = \pi_{32}\pi_{43}$, $\cdots$. Puisque les $\Gamma_i^k (1 \leqq i \leqq m)$ sont des (k, R)-cycles mod $(AB - U)$ dans AB, ce sont aussi des (k, R)-cycles mod $(A - U)$ dans A. Or $\beta_k(V, U; A) = 0$, de manière que $\Gamma_i^k \sim 0 \bmod (A - V)$ dans A. Pareillement $\Gamma_i^k \sim 0 \bmod (B - V)$ dans B, car $\beta_k(V, U; B) = 0$. Donc il existe des $(k + 1, \mathfrak{U}_4)$-chaînes $C_{i1}^{k+1}(\mathfrak{U}_4) \subset A$ et $C_{i2}^{k+1}(\mathfrak{U}_4) \subset B$ telles que

(1) $C_{i1}^{k+1}(\mathfrak{U}_4) \to \Gamma_i^k(\mathfrak{U}_4) \bmod (A - V)$, $\quad C_{i2}^{k+1}(\mathfrak{U}_4) \to \Gamma_i^k(\mathfrak{U}_4) \bmod (B - V)$.

Posons $C_i^{k+1}(\mathfrak{U}_4) = C_{i1}^{k+1}(\mathfrak{U}_4) - C_{i2}^{k+1}(\mathfrak{U}_4)$. Alors $C_i^{k+1}(\mathfrak{U}_4) \to 0 \bmod (R - V)$ de sorte que, d'après la définition de $\mathfrak{U}_4$, $\pi_{43} C_i^{k+1}(\mathfrak{U}_4)$ est un $(k + 1, \mathfrak{U}_3)$-cycle essentiel mod $(R - V)$.

Il ne reste qu'à démontrer que les $\pi_{43} C_i^{k+1}(\mathfrak{U}_4)(1 \leqq i \leqq m)$ ne sont liés par aucune homologie mod $(R - W)$, car alors (v. *Homologie*, II, 28) on a $\beta_{k+1}(W, V; R) \geqq m$, c.q.f.d.

Soit donc $\pi_{43} \sum_{i=1}^{m} r_i C_i^{k+1}(\mathfrak{U}_4) \sim 0 \bmod (R - W)(r_i \in \mathfrak{R})$; il s'agit d'en déduire que $r_1 = \cdots = r_m = 0$. L'homologie qui vient d'être écrite signifie qu'il existe une $(k + 2, \mathfrak{U}_3)$-chaîne $D^{k+2}(\mathfrak{U}_3)$ telle que

(2) $D^{k+2}(\mathfrak{U}_3) \to \pi_{43} \sum_{i=1}^{m} r_i C_i^{k+1}(\mathfrak{U}_4) \bmod (R - W)$.

Puisque $R = A + B$, on peut poser $D^{k+2}(\mathfrak{U}_3) = D_1^{k+2}(\mathfrak{U}_3) - D_2^{k+2}(\mathfrak{U}_3)$, où $D_2^{k+2}(\mathfrak{U}_3) \subset A$, $D_2^{k+2}(\mathfrak{U}_3) \subset B$. Il résulte de (1) et (2) que

$$(3)\quad \pi_{43}\sum_{i=1}^{m} r_i C_{i1}^{k+1}(\mathfrak{U}_4) - FD_1^{k+2}(\mathfrak{U}_4) = \pi_{43}\sum_{i=1}^{m} r_i C_{i2}^{k+1}(\mathfrak{U}_4) - FD_2^{k+2}(\mathfrak{U}_3)$$

mod $(R - W)$. Désignons par $E^{k+1}(\mathfrak{U}_3)$ cette partie d'un (et donc aussi de l'autre) membre de (3) dont les simplexes sont $\not\equiv 0$ mod $(R - W)$. Alors $E^{k+1}(\mathfrak{U}_3) \subset A$, $E^{k+1}(\mathfrak{U}_3) \subset B$, d'où $E^{k+1}(\mathfrak{U}_3) \subset AB$ d'après la définition de $\mathfrak{U}_3$. En outre

$$FD_1^{k+2}(\mathfrak{U}_3) = \pi_{43}\sum_{i=1}^{m} r_i C_{i1}^{k+1}(\mathfrak{U}_4) - E^{k+1}(\mathfrak{U}_3) \mod (R - W),$$

d'où $\pi_{43}\sum_{i=1}^{m} r_i C_{i1}^{k+1}(\mathfrak{U}_4) - E^{k+1}(\mathfrak{U}_3) \to 0$ mod $(R - W)$. En tenant compte de (1), on en déduit que $FE^{k+1}(\mathfrak{U}_3) - \pi_{43}\sum_{i=1}^{m} r_i \Gamma_i^k(\mathfrak{U}_4) \subset R - W$. Donc la $(k, \mathfrak{U}_2)$-chaîne

$$(4)\quad F\pi_{32}E^{k+1}(\mathfrak{U}_3) - \pi_{42}\sum_{i=1}^{m} r_i \Gamma_i^k(\mathfrak{U}_4)$$

est située dans $(R - W)$. Puisque $E^{k+1}(\mathfrak{U}_3) \subset AB$, $\Gamma_i^k(\mathfrak{U}_4) \subset AB$, la chaîne (4) est aussi située dans AB. D'après la définition de $\mathfrak{U}_2$, la chaîne (4) est donc située dans $AB - W$. Donc $\pi_{32}E^{k+1}(\mathfrak{U}_3) \to \pi_{42}\sum_{i=1}^{m} r_i \Gamma_i^k(\mathfrak{U}_4)$ mod $(AB - W)$, d'où

$$\pi_{31}E^{k+1}(\mathfrak{U}_3) \to \pi_{41}\sum_{i=1}^{m} r_i \Gamma_i^k(\mathfrak{U}_4) \mod (AB - W).$$

Donc $\pi_{41}\sum_{i=1}^{m} r_i \Gamma_i^k(\mathfrak{U}_4) \sim 0$ mod $(AB - W)$ dans AB. Or, Γ_i^k étant un (k, R)-cycle mod $(AB - U)$ dans AB, on a $\pi_{41}\Gamma_i^k(\mathfrak{U}_4) \sim \Gamma_i^k(\mathfrak{U}_1)$ mod $(AB - U)$ dans AB. Comme $W \subset U$, il en résulte que $\sum_{i=1}^{m} r_i \Gamma_i^k(\mathfrak{U}_1) \sim 0$ mod $(AB - W)$ dans AB. D'après la définition de $\mathfrak{U}_1$, on en déduit que $r_1 = \cdots = r_m = 0$, c.q.f.d.

III.

7. Soit $n = 1, 2, 3, \cdots$. Soit R un espace topologique. Soit $S \subset R$. Nous dirons que R est *une variété à n dimensions* mod S *d'ordre* 0, ou *une* $V_0^n(S)$, si les axiomes suivants sont vérifiés:

1°. *R est un espace bicompact.*[12]

[12] Cela signifie: 1° a et b étant deux points distincts de R, il existe des ensembles ouverts U et V tels que $a \in U$, $b \in V$, $UV = 0$; 2° de chaque famille d'ensembles ouverts recouvrant R on peut extraire une famille finie recouvrant R.

2°. *Chaque sous-ensemble ouvert de R est un F_σ dans R.*

3°. *L'ensemble S est fermé dans R.*

4°. $\dim R - S = n$.

5°. *R est localement connexe d'ordre n rel. à $\mathfrak{R}$ à chaque point $x \in R - S$.*[13]

6°. *Pour $x \in R - S$ on a $\beta_n(x, R) = 1$.*

7°. *A étant un sous-ensemble fermé de R, on a $\beta_n(x, A) = 0$ pour chaque $x \in A$. $\overline{R - A} - S$.*

8. *Soit $S \subset T = \overline{T} \subset R$. Si R est une $V_0^n(S)$, R est* évidemment *aussi une $V_0^n(T)$.*

9. *Soit R un espace topologique; soit $S \subset R$, R est une $V_0^n(S)$ si et seulement s'il satisfait aux axiomes A_1, A_2, A_3, A_4, B, D_1, D_2, E énoncés dans Variétés, nos.* 1, 2, 3, 7, 9, 11, 12, 13.

Démonstration. Les axiomes A_1, A_2, A_3, A_4, B disent la même chose que les axiomes 1° — 5°. On doit donc déduire d'abord la validité de D_1, D_2, et E en supposant 1° — 7°, ce qui sera fait aux n°s 9.11 (pour D_1), 9.12 (pour D_2) et 9.13 (pour E), et ensuite la validité de 6° et 7° en supposant A_1, A_2, A_3, A_4, B, D_1, D_2, E, ce qui sera fait aux n°s 9.21 (pour 6°) et 9.22 (pour 7°).

9.11 Soit Γ^{n-1} un $(n-1, R)$-cycle du type t_2 dans A (*Variétés*, 10). Soit H le porteur de l'homologie $\Gamma^{n-1} \sim 0$ (*Variétés* 10.3) de manière que $H = \overline{H} \subset R - S$. On doit prouver (*Variétés*, 11) que l'ensemble $H - A$ est ouvert. Supposons le contraire. Il existe donc un point $a \in (H - A)\,\overline{R - H} \subset H \cdot \overline{R - H} - S$. D'après l'axiome 7° on a $\beta_n(a, H) = 0$. Or soit C^n le cycle relatif correspondant à Γ^{n-1} d'après *Variétés* 10.1. Donc C^n est un (n, R)-cycle mod A dans H et $FC^n(\mathfrak{U}) \sim \Gamma^{n-1}(\mathfrak{U})$ dans A pour chaque réseau $\mathfrak{U}$. Soit U un entourage de a si petit que $UA = 0$ Purisque $\beta_n(a, H) = 0$, on a $\beta_n(a, U; H) = 0$. Donc il existe un entourage $V \subset U$ de a tel que $\beta_n(V, U; H) = 0$. Or $A \subset H - U$, de sorte que C^n est un (n, R)-cycle mod $(H - U)$ dans H. Comme $\beta_n(V, U; H) = 0$, on a $C^n \sim 0$ mod $(H - V)$ dans H. Donc il existe pour chaque réseau $\mathfrak{U}$ une $(n+1, \mathfrak{U})$-chaîne $M^{n+1}(\mathfrak{U}) \subset H$ et une $(n, \mathfrak{U})$-chaîne $D^n(\mathfrak{U}) \subset H - V$ telles que $M^{n+1}(\mathfrak{U}) \to C^n(\mathfrak{U}) - D^n(\mathfrak{U}) \to 0$, d'où $FD^n(\mathfrak{U}) \sim \Gamma^{n-1}(\mathfrak{U})$ dans A. Or ceci signifie qu'il existe pour chaque $\mathfrak{U}$ une $(n, \mathfrak{U})$-chaîne $E^n(\mathfrak{U}) \subset A$ telle que $D^n(\mathfrak{U}) - E^n(\mathfrak{U}) \to \Gamma^{n-1}(\mathfrak{U})$, d'où $\Gamma^{n-1} \sim 0$ dans $(H - U) + A \subset H$. D'après la définition de H il en résulte que $a \in H - U$, ce qui est une contradiction.

9.12. Soit $a \in R - S$. On doit prouver (*Variétés*, 12) que le point a est situé à l'intérieur d'un $(n-1, R)$-cycle du type t_2. Soit P_2 un sommet du réseau gén. $\mathfrak{P}_2$ (*Variétés*, 9.2) contenant le point a. D'après l'axiome 6° on a $\beta_n(a, R) = 1$ d'où, comme on le voit sans peine, $\beta_n(a, \overline{P}_2) = 1$. Donc il existe un entourage $U \subset P_2$ de a tel que $\beta_n(a, U; \overline{P}_2) = 1$. Donc il existe un entourage $V \subset U$ de a tel que $\beta_n(W, U; \overline{P}_2) = 1$ pour chaque entourage $W \subset V$ de a.

[13] Cela signifie (v. le n°18 du Mémoire cité au n° 4): Chaque entourage P de x contient un entourage Q de x tel que $\Gamma^n \sim 0$ dans $\overline{P}$ pour chaque (n, R)-cycle absolu Γ^n dans $\overline{Q}$.

En particulier $\beta_n(V, U; \overline{P}_2) = 1$ de sorte qu'il existe un (n, R)-cycle C^n mod $(\overline{P}_2 - U)$ dans $\overline{P}_2$ qui n'est pas ~ 0 mod $(\overline{P}_2 - V)$ dans $\overline{P}_2$, tandis que chaque (n, R)-cycle mod $(\overline{P}_2 - U)$ dans $\overline{P}_2$ est $\sim rC^n (r \in \Re)$ mod $(\overline{P}_2 - V)$ dans $\overline{P}_2$ et par suite aussi mod $(\overline{P}_2 - W)$ dans $\overline{P}_2$ pour chaque entourage $W \subset V$ de a. Comme $\beta_n(W, U; \overline{P}_2) = 1$, il en résulte que C^n n'est ~ 0 mod $(\overline{P}_2 - W)$ dans P pour aucun choix de $W \subset V$. Pour chaque réseau $\mathfrak{U}$ soit $\Gamma^{n-1}(\mathfrak{U}) = FC^n(\mathfrak{U})$. C^n étant un (n, R)-cycle mod $(\overline{P}_2 - U)$ dans $\overline{P}_2$, on voit sans peine que Γ^{n-1} est un $(n-1, R)$-cycle absolu dans $(\overline{P}_2 - U)$. Evidemment $\Gamma^{n-1} \sim 0$ dans $\overline{P}_2$; donc (*Variétés*, 10) Γ^{n-1} est un $(n-1, R)$-cycle du type t_2 dans $A = \overline{P}_2 - U$. Soit H le porteur de l'homologie $\Gamma^{n-1} \sim 0$. On doit prouver que $a \in H - A$. Or $\Gamma^{n-1} \sim 0$ dans $\overline{P}_2$, de sorte que (*Variétés* 10.3)$H \subset \overline{P}_2$. Supposons par impossible que le point a n'appartienne pas à $H - A$. Alors il existe un entourage $W \subset V$ de a tel que $WH = 0$, d'où $H \subset \overline{P}_2 - W$. Or $\Gamma^{n-1} \sim 0$ dans H d'après la définition de H, de sorte que pour chaque réseau $\mathfrak{U}$ il existe une $(n, \mathfrak{U})$-chaîne $D^n(\mathfrak{U}) \subset H$ telle que $FD^n(\mathfrak{U}) \sim \Gamma^{n-1}(\mathfrak{U})$ dans $A = \overline{P}_2 - U$, d'où $F[C^n(\mathfrak{U}) - D^n(\mathfrak{U})] \sim 0$ dans $\overline{P}_2 - U$. Donc il existe une $(n, \mathfrak{U})$-chaîne $E^n(\mathfrak{U}) \subset \overline{P}_2 - U$ telle que $C^n(\mathfrak{U}) - D^n(\mathfrak{U}) - E^n(\mathfrak{U}) \to 0$. D'après *Homologie*, II, 21 on peut s'arranger de façon que $\{C^n(\mathfrak{U}) - D^n(\mathfrak{U}) - E^n(\mathfrak{U})\}$ soit un (n, R)-cycle dans $\overline{P}_2$. Or $P_2 \in \mathfrak{P}_2$, de sorte que $C^n(\mathfrak{U}) - D^n(\mathfrak{U}) - E^n(\mathfrak{U}) \sim 0$ dans $\overline{P}_2$ d'après *Variétés*, 9.1–9.3. Donc il existe pour chaque $\mathfrak{U}$ une $(n+1, \mathfrak{U})$-chaîne $M^{n+1}(\mathfrak{U}) \subset \overline{P}_2$ telle que $C^n(\mathfrak{U}) = FM^{n+1}(\mathfrak{U}) + D^n(\mathfrak{U}) + E^n(\mathfrak{U})$. Or $D^n(\mathfrak{U}) \subset H \subset \overline{P}_2 - W$, $E^n(\mathfrak{U}) \subset \overline{P}_2 - U \subset \overline{P}_2 - W$, de sorte que $C^n \sim 0$ mod $(\overline{P}_2 - W)$ dans $\overline{P}_2$, ce qui est une contradiction.

9.13. Soit $a \in R - S$; soit Q un entourage de a. D'après l'axiome 6° on a $\beta_n(a, R) = 1$ et par suite aussi $\beta_n(a, \bar{Q}) = 1$. Donc il existe un entourage $Q_1 \subset Q$ de a tel que $\beta_n(a, Q_2; \bar{Q}) = 1$ pour chaque entourage $Q_2 \subset Q_1$ de a. On peut supposer que Q_1 fasse partie d'un sommet du réseau gén. $\mathfrak{P}_2$ (*Variétés*, 9.3). Soit Q_2 un entourage de a contenu dans Q_1. Comme $\beta_n(a, Q_2; \bar{Q}) = 1$, il existe un entourage Q_3 de a tel que $\beta_n(Q_3, Q_2; \bar{Q}) = 1$.

Soient Γ_1^{n-1}, Γ_2^{n-1} deux $(n-1, R)$-cycles absolus dans $(\bar{Q}_1 - Q_2)$ qui soient ~ 0 dans $\bar{Q}_1$. Il suffit de prouver (*Variétés*, 13) qu'il existe deux nombres $r_1, r_2 \in \Re$, dont un au moins $\neq 0$, tels que $r_1\Gamma_1^{n-1} + r_2\Gamma_2^{n-1} \sim 0$ dans $(\bar{Q} - Q_3)$. Soient C_1^n, C_2^n les cycles relatifs correspondant a Γ_1^{n-1}, Γ_2^{n-1} d'après *Variétés*, 10.1. Ce sont donc des (n, R)-cycles mod $(\bar{Q}_1 - Q_2)$ dans $\bar{Q}_1$ tels que $FC_i^n(\mathfrak{U}) \sim \Gamma_i^{n-1}(\mathfrak{U})$ dans $(\bar{Q}_1 - Q_2)$ pour $i = 1, 2$ et pour chaque réseau $\mathfrak{U}$. Les C_i^n étant des (n, R)-cycles mod $(\bar{Q}_1 - Q_2)$ dans $\bar{Q}_1$, l'équation $\beta_n(Q_3, Q_2; \bar{Q}) = 1$ entraîne qu'il existe deux nombres $r_1, r_2 \in \Re$, dont un au moins $\neq 0$, tels que $r_1C_1^n + r_2C_2^n \sim 0$ mod $(\bar{Q} - Q_3)$ dans $\bar{Q}$. Donc il existe pour chaque réseau $\mathfrak{U}$ une $(n+1, \mathfrak{U})$-chaîne $M^{n+1}(\mathfrak{U}) \subset \bar{Q}$ et une $(n, \mathfrak{U})$-chaîne $D^n(\mathfrak{U}) \subset \bar{Q} - Q_3$ telles que

$$M^{n+1}(\mathfrak{U}) \to r_1C_1^n(\mathfrak{U}) + r_2C_2^n(\mathfrak{U}) - D^n(\mathfrak{U}) \to 0.$$

Or $FC_i^n(\mathfrak{U}) \sim \Gamma_i^{n-1}(\mathfrak{U})$ dans $\bar{Q}_1 - Q_2 \subset \bar{Q} - Q_3$, de sorte qu'il existe deux $(n, \mathfrak{U})$-chaînes $E_i^n(\mathfrak{U}) \subset \bar{Q} - Q_3$ $(i = 1, 2,)$ telles que $FC_i^n(\mathfrak{U}) = \Gamma_i^{n-1}(\mathfrak{U}) + FE_i^n(\mathfrak{U})$,

d'où $r_1\Gamma_1^{n-1}(\mathfrak{U}) + r_2\Gamma_2^{n-1}(\mathfrak{U}) = F[D^n(\mathfrak{U}) - r_1E_1^n(\mathfrak{U}) - r_2E_2^n(\mathfrak{U})]$, donc $r_1\Gamma_1^{n-1} + r_2\Gamma_2^{n-1} \sim 0$ dans $\bar{Q} - Q_3$, c.q.f.d.

9.21. Soit $a \in R - S$. On doit prouver que $\beta_n(a, R) = 1$. Si l'on avait $\beta_n(a, R) > 1$, il existerait un entourage U de a tel que $\beta_n(V, U; R) > 1$ pour chaque entourage $V \subset U$ de a. Or c'est impossible d'après *Variétés*, 14.1. Donc il suffit de prouver que $\beta_n(a, R) \geqq 1$.

D'après l'axiome D_2 il existe un $(n-1, R)$-cycle Γ^{n-1} du type t_2 dans A tel que le point a appartient à l'intérieur de Γ^{n-1}. Soit H le porteur de l'homologie $\Gamma^{n-1} \sim 0$ de sorte que $a \in H - A$. D'après l'axiome D_1, l'ensemble $H - A$ est un entourage de a. Evidemment $\beta_n(a, R) = \beta_n(a, H)$. Il suffit donc de prouver que $\beta_n(a, H) \geqq 1$. Supposons au contraire que $\beta_n(a, H) = 0$, d'où $\beta_n(a, H - A; H) = 0$. Donc il existe un entourage $V \subset H - A$ de a tel que $\beta_n(V, H - A; H) = 0$. Soit C^n le cycle relatif correspondant à Γ^{n-1} d'après *Variétés*, 10.1. Donc C^n est un (n, R)-cycle mod A dans H tel que $FC^n(\mathfrak{U}) \sim \Gamma^{n-1}(\mathfrak{U})$ dans A pour chaque réseau $\mathfrak{U}$. Comme $\beta_n(V, H - A; H) = 0$, on a $C^n \sim 0$ mod $(H - V)$ dans H. Donc il existe pour chaque $\mathfrak{U}$ une $(n+1, \mathfrak{U})$-chaîne $M^{n+1}(\mathfrak{U}) \subset H$ et une $(n, \mathfrak{U})$-chaîne $D^n(\mathfrak{U}) \subset H - V$ telles que $M^{n+1}(\mathfrak{U}) \to C^n(\mathfrak{U}) - D^n(\mathfrak{U}) \to 0$. Or $FC^n(\mathfrak{U}) \sim \Gamma^{n-1}(\mathfrak{U})$ dans A de sorte qu'il existe une $(n, \mathfrak{U})$-chaîne $E^n(\mathfrak{U}) \subset A$ telle que $\Gamma^{n-1}(\mathfrak{U}) = F[C^n(\mathfrak{U}) - E^n(\mathfrak{U})]$. Donc $D^n(\mathfrak{U}) + E^n(\mathfrak{U}) \to \Gamma^{n-1}(\mathfrak{U})$, d'où $\Gamma^{n-1} \sim 0$ dans $A + (H - U)$, donc $H = A + (H - U)$ d'après la définition de H, d'où la contradiction $a \in A + (H - U)$.

9.22. Soit A un sous-ensemble fermé de R et soit $a \in A \cdot \overline{R - A} - S$. On doit prouver que $\beta_n(a, A) = 0$. Supposons au contraire que $\beta_n(a, A) \geqq 1$. Il en résulte que $\beta_n(a, U; A) \geqq 1$ pour chaque entourage U de a suffisamment petit. D'après *Variétés* 12.1 on peut choisir cet entourage U de manière qu'il existe un $(n-1, R)$-cycle absolu Δ^{n-1} dans $\overline{U} - U$ tel que $\Delta^{n-1} \sim 0$ dans $\overline{U}$, mais non pas $\Delta^{n-1} \sim 0$ dans un vrai sous-ensemble fermé de $\overline{U}$. D'après *Variétés*, 14.1 il existe un entourage $V \subset U$ de a tel qu'à chaque couple C^n, D^n de (n, R)-cycles mod $(R - U)$ on puisse attacher deux nombres $r, s \in \mathfrak{R}$ ($r \neq 0$ ou $s \neq 0$) de manière que $rC^n + sD^n \sim 0$ mod $(R - V)$. On peut supposer que $\bar{V} \subset R - S$. Soit (*Variétés*, 1.2) W un entourage de a tel que $\overline{W} \subset V$. Comme $\beta_n(a, U; A) \geqq 1$, on a $\beta_n(W, U; A) \geqq 1$. Comme $a \in \overline{R - A}$, il existe un ensemble ouvert $Q \neq 0$ tel que $A\bar{Q} = 0$, $\bar{Q} \subset V$. Soit Φ la famille de tous les réseaux $\mathfrak{U}$ d'ordre mod $S \leqq n$ (*Variétés*, 7.2) et tels que 1° $u \in \mathfrak{U}$, $uS \neq 0$ entraîne que $u \subset R - V$; 2° $u \in \mathfrak{U}$, $u - V \neq 0$ entraîne que $u(Q + W) = 0$, 3° $u \in \mathfrak{U}$, $uA \neq 0$ entraîne que $uQ = 0$. On voit sans peine (v. *Variétés*, 7.4) que la famille Φ est complète (rel. à la famille fondamentale de tous les réseaux ouverts; v. *Homologie*, II, 30 et III, 2). Comme $\beta_n(W, U; A) \geqq 1$, il existe un (n, R)-cycle C^n mod $(A - U)$ dans A qui n'est pas ~ 0 mod $(A - W)$ dans A. Comme $\Delta^{n-1} \sim 0$ dans $\overline{U}$, il existe pour chaque réseau $\mathfrak{U}$ une $(n, \mathfrak{U})$-chaîne $D^n(\mathfrak{U}) \subset U$ telle que $FD^n(\mathfrak{U}) \sim \Delta^{n-1}(\mathfrak{U})$ dans $\overline{U} - U$. D'après *Homologie*, II, 21 on peut s'arranger de façon que $D^n = \{D^n(\mathfrak{U})\}$ soit un (n, R)-cycle mod $(\overline{U} - U)$ dans $\overline{U}$. Donc C^n et D^n sont deux (n, R)-cycles mod $(R - U)$ de

manière qu'il existe deux nembres $r, s \in \mathfrak{R}$ ($r \neq 0$ ou $s \neq 0$) tels que $rC^n + sD^n \sim 0$ mod $(R - V)$. Donc il existe pour chaque réseau $\mathfrak{U}$ une $(n+1, \mathfrak{U})$-chaîne $M^{n+1}(\mathfrak{U})$ et une $(n, \mathfrak{U})$-chaîne $E^n(\mathfrak{U}) \subset R - V$ telles que $FM^{n+1}(\mathfrak{U}) = rC^n(\mathfrak{U}) + sD^n(\mathfrak{U}) - E^n(\mathfrak{U})$. Soit $\mathfrak{U} \in \Phi$. L'ordre mod S de $\mathfrak{U}$ étant $\leqq n$, chaque simplexe de $M^{n+1}(\mathfrak{U})$ contient un sommet u tel que $uS \neq 0$, d'où $u \subset R - V$; il en résulte que $M^{n+1}(\mathfrak{U}) \subset R - V$. Donc $rC^n(\mathfrak{U}) + sD^n(\mathfrak{U}) \subset R - V$ pour chaque $\mathfrak{U} \in \Phi$.

Soit d'abord $s = 0$, d'où $r \neq 0$ et par suite $C^n(\mathfrak{U}) \subset R - V$ pour $\mathfrak{U} \in \Phi$. Comme $C^n(\mathfrak{U}) \subset A$ et comme aucun sommet de $\mathfrak{U} \in \Phi$ ne peut rencontrer simultanément W et $R - V$, on a $C^n(\mathfrak{U}) \subset A - W$ pour $\mathfrak{U} \in \Phi$. La famille Φ étant complète, on arrive à la contradiction que $C^n \sim 0$ mod $(A - W)$ dans A.

Passons au cas $s \neq 0$. Pour $\mathfrak{U} \in \Phi$ on a $C^n(\mathfrak{U}) \subset A$, $rC^n(\mathfrak{U}) + sD^n(\mathfrak{U}) \subset R - V$, $s \neq 0$, d'autre part un sommet de $\mathfrak{U}$ ne peut rencontrer simultanément ni A et Q, ni $R - V$ et Q. Il en résulte que pour $\mathfrak{U} \in \Phi$ aucun sommet d'un simplexe de $D^n(\mathfrak{U})$ ne peut rencontrer Q. Comme $D^n(\mathfrak{U}) \subset \overline{U}$, on a $D^n(\mathfrak{U}) \subset \overline{U} - Q$ pour $\mathfrak{U} \in \Phi$. Or $FD^n(\mathfrak{U}) \sim \Delta^{n-1}(\mathfrak{U})$ dans $\overline{U} - U \subset \overline{U} - Q$ de manière que $\Delta^{n-1}(\mathfrak{U}) \sim 0$ dans $(\overline{U} - Q)$ pour $\mathfrak{U} \in \Phi$. La famille Φ étant complète, on arrive à la contradiction que $\Delta^{n-1} \sim 0$ dans $(\overline{U} - Q)$.

10. Soit R une $V_0^n(S)$. On dit que R est *orientable* mod S s'il existe un (n, R)-cycle G^n mod S tel que, pour $A = \bar{A} \subset R$, $R - (A + S) \neq 0$, G^n n'est $\sim$ mod S à aucun (n, R)-cycle mod S dans $(A + S)$. Un tel cycle G^n s'appelle alors un *(n, R)-cycle principal* mod S. *Orienter* R mod S signifie que l'on choisit un (n, R)-cycle principal mod S bien déterminé.

11. *G^n est un (n, R)-cycle principal* mod S *si et seulement si pour* $A = \bar{A} \subset R$, $R - (A + S) \neq 0$ *on n'a jamais* $G^n \sim 0$ mod $(A + S)$.

Démonstration. I. Supposons que $G^n \sim 0$ mod $(A + S)$. On doit prouver que $G^n \sim H^n$ mod S, H^n étant un (n, R)-cycle mod S dans $(A + S)$. Or pour chaque réseau $\mathfrak{U}$ il existe une $(n+1, \mathfrak{U})$-chaîne $M^{n+1}(\mathfrak{U})$ et une $(n, \mathfrak{U})$-chaîne $H^n(\mathfrak{U}) \subset A + S$ telles que $M^{n+1}(\mathfrak{U}) \to G^n(\mathfrak{U}) - H^n(\mathfrak{U})$ mod S, d'où $H^n(\mathfrak{U}) \to 0$ mod S, car $G^n(\mathfrak{U}) \to 0$ mod S. D'après *Homologie*, II, 21 on peut s'arranger de façon que $H^n = \{H^n(\mathfrak{U})\}$ soit un (n, R)-cycle mod S dans $(A + S)$. Evidemment $G^n \sim H^n$ mod S.

II. Supposons que $G^n \sim H^n$ mod S, H^n étant un (n, R)-cycle mod S dans $(A + S)$. Alors évidemment $G^n \sim 0$ mod $(A + S)$.

12. *Soit* $S \subset T = \overline{T} \subset R$. Evidemment, *si R est une $V_0^n(S)$ orientable* mod S, *R est aussi une $V_0^n(T)$ orientable* mod T.

13. *Soit R une $V_0^n(S)$. R est orientable* mod S *si et seulement si l'axiome F (Variétés*, 16) *est satisfait.*

Démonstration. I. Supposons que l'axiome F soit vérifié. Nous avons construit un (n, R)-cycle principal mod S dans *Variétés* 17–17.5.

Soit G^n un (n, R)-cycle principal mod S. Pour chaque $x \in R - S$, soit (v. *Variétés*, 15) $\Theta_2(x)$ la famille de tous les $(n-1, R)$-cycles du type t_2 dans A, A étant assujetti à la condition de ne pas contenir x. On doit (*Variétés*, 15 et 16) attacher à chaque $a \in R - S$, $\Gamma^{n-1} \in \theta_2(a)$ un nombre $\omega(a, \Gamma^{n-1}) \in \mathfrak{R}$ selon les conditions suivantes: $1°\omega(a, \Gamma^{n-1}) = 0$ si et seulement si le porteur de l'homologie $\Gamma^{n-1} \sim 0$ ne contient pas le point a, $2°$ pour $\Gamma^{n-1} \in \theta_2(a)$, $r \in \mathfrak{R}$ on doit avoir $\omega(a, r\,\Gamma^{n-1}) = r\,\omega(a, \Gamma^{n-1})$; $3°$ pour Γ_1^{n-1}, Γ_2^{n-2}, $\Gamma_1^{n-1} + \Gamma_2^{n-1} \in \theta_2(a)$ on doit avoir $\omega(a, \Gamma_1^{n-1} + \Gamma_2^{n-1}) = \omega(a, \Gamma_1^{n-1}) + \omega(a, \Gamma_2^{n-1})$, $4°$ $a \in R - S$ et $\Gamma^{n-1} \in \theta_2(a)$ étant donnés, il doit exister un entourage V de a tel que $x \in V$ entraîne que $\Gamma^{n-1} \in \theta_2(x)$ et que $\omega(x, \Gamma^{n-1}) = \omega(a, \Gamma^{n-1})$.

Soit donc $a \in R - S$ et soit $\Gamma^{n-1} \sim \theta_2(a)$ un $(n-1, R)$-cycle du type t_2 dans A. Soit H le porteur de l'homologie $\Gamma^{n-1} \sim 0$. Lorsque $a \in R - H$, soit $\omega(a, \Gamma^{n-1}) = 0$. Supposons donc que $a \in H$. Soit C^n le (n, R)-cycle mod A dans H déduit de Γ^{n-1} d'après *Variétés*, 10.1. D'après *Variétés*, 11, l'ensemble $H - A$ est un entourage de a. Evidemment C^n et G^n sont deux (n, R)-cycles mod $R - (H - A)$ dans R. D'après *Variétés*, 14.1 il existe un entourage $V \subset H - A$ de a et deux nombres $r, s \in \mathfrak{R}$ ($r \neq 0$ ou $s \neq 0$) tels que $rC^n + sG^n \sim 0$ mod $(R - V)$. D'après 11 on a $r \neq 0$. Posons $\omega(a, \Gamma^{n-1}) = -s/r$.

On vérifie sans peine que les propriétés $1° - 4°$ sont vérifiées.

14. *Soit R un $V_0^n(S)$. Soit $a \in R - S$. Il existe un entourage $V \subset R - S$ de a tel que $R - S$ est orientable* mod $(R - V)$.

Démonstration. D'après 8,6° on a $\beta_n(a, R) = 1$. Donc il existe un entourage $U \subset R - S$ de a tel que $\beta_n(a, U; R) = 1$. Donc il existe un entourage $V \subset U$ de a tel que $\beta_n(V, U; R) = 1$. Donc il existe un (n, R)-cycle G^n mod $(R - U)$ qui n'est pas ~ 0 mod $(R - V)$. Soit W un ensemble ouvert tel que $0 \neq W \subset V$. Soit $b \in W$. On voit sans peine que

$$1 \leqq \beta_n(b, U; R) \leqq \beta_n(W, U; R) \leqq \beta_n(V, U; R) = 1,$$

d'où $\beta_n(W, U; R) = 1$. On en déduit sans peine en premier lieu qu'il existe un (n, R)-cycle C^n mod $(R - U)$ qui n'est pas ~ 0 mod $(R - W)$ et en second lieu qu'il existe un nombre $r \in \mathfrak{R}$ tel que $G^n \sim rC^n$ mod $(R - W)$ et donc aussi mod $(R - V)$. Puisque G^n n'est pas ~ 0 mod $(R - V)$, on a $r \neq 0$. Donc G^n n'est ~ 0 mod $(R - W)$ pour aucun choix de l'ensemble W ouvert et tel que $0 \neq W \subset V$. D'après 11 G^n est un (n, R)-cycle principal mod $(R - V)$.

15. *Soit R une $V_0^n(S)$. Pour $0 \leqq k \leqq n - 1$ nous allons considérer les deux axiomes suivants:*

I^k. *R est localement connexe d'ordre k rel. à $\mathfrak{R}$ à chaque point $x \in R - S$.*

II^k. *Pour $x \in R - S$ on a $\beta_k(x, R) = 0$.*

On dit que R est une *variété à n dimensions d'ordre p* $(1 \leqq p \leqq n)$ mod S, ou une $V_p^n(S)$, si c'est une $V_0^n(S)$ vérifiant les axiomes I^k et II^k pour $n - p \leqq k \leqq n - 1$.

16. *Soit* $S \subset T = \overline{T} \subset R$, $1 \leqq p \leqq n$. *Si* R *est une* $V_p^n(S)$, R *est* évidemment *aussi une* $V_p^n(T)$.

17. *Soit* $0 \leqq p \leqq q \leqq n$. Evidemment *une* $V_q^n(S)$ *est aussi une* $V_p^n(S)$.

18. *Soit* $0 \leqq k \leqq n-1$. *Les deux axiomes* I^k *et* II^k *sont équivalents aux axiomes* G^k (*Variétés*, 21) et H^{k-1} (*Variétés* 27; pour $k = 0$ l'axiome H^{-1} doit signifier que l'ensemble $R - S$ est dense en soi).[14]

DÉMONSTRATION. L'axiome I^k est identique à l'axiome G^k. Soit d'abord $k = 0$. On doit prouver que si l'ensemble $R - S$ est localement connexe (au sens classique), l'axiome II^0 est équivalent à l'axiome H^{-1}. Or c'est une conséquence immédiate du théorème du n°3.

Passons au cas $1 \leqq k \leqq n-1$. Il suffit de déduire d'abord l'axiome H^{k-1} de l'axiome II^k, ce qui sera fait au n°18.1, et ensuite l'axiome II^k des axiomes G^k et H^{k-1}, ce qui sera fait au n°18.2.

18.1. Supposons la validité de II^k. Soit $a \in R - S$ et soit P un entourage donné de a. On a $\beta_k(a, R) = 0$, d'où $\beta_k(a, \overline{P}) = 0$. Soient P_1 et P_2 deux entourages de a tels que $P_2 \subset P_1 \subset P$. Puisque $\beta_k(a, \overline{P}) = 0$, on a $\beta_k(a, P_2; \overline{P}) = 0$. Donc il existe un entourage $P_3 \subset P_2$ de a tel que $\beta_k(P_3, P_2; \overline{P}) = 0$. Soit Γ^{k-1} un $(k-1, R)$-cycle absolu dans $(\overline{P}_1 - P_2)$; soit $\Gamma^{k-1} \sim 0$ dans $\overline{P}_1$. Il suffit (*Variétés* 21) d'en déduire que $\Gamma^{k-1} \sim 0$ dans $(\overline{P} - P_3)$.

Comme $\Gamma^{k-1} \sim 0$ dans $\overline{P}_1$, il existe pour chaque réseau $\mathfrak{U}$ une $(k, \mathfrak{U})$-chaîne $C^k(\mathfrak{U}) \subset \overline{P}_1$ telle que $FC^k(\mathfrak{U}) \sim \Gamma^{k-1}(\mathfrak{U})$ dans $\overline{P}_1 - P_2$. D'après *Homologie*, II, 21 on peut s'arranger de façon que $C^k = \{C^k(\mathfrak{U})\}$ soit un (k, R)-cycle mod $(\overline{P}_1 - P_2)$ dans $\overline{P}_1$. Comme $P_1 \subset P$, $\beta_k(P_3, P_2; \overline{P}) = 0$, on a $C^k \sim 0$ mod $(\overline{P} - P_3)$ dans $\overline{P}$. Donc il existe pour chaque réseau $\mathfrak{U}$ une $(k+1, \mathfrak{U})$-chaîne $M^{k+1}(\mathfrak{U}) \subset \overline{P}$ et une $(k, \mathfrak{U})$-chaîne $D^k(\mathfrak{U}) \subset \overline{P} - P_3$ telles que $M^{k+1}(\mathfrak{U}) \to C^k(\mathfrak{U}) - D^k(\mathfrak{U}) \to 0$. Comme $FC^k(\mathfrak{U}) \sim \Gamma^{k-1}(\mathfrak{U})$ dans $(\overline{P}_1 - P_2)$, il existe une $(k, \mathfrak{U})$-chaîne $E^k(\mathfrak{U}) \subset \overline{P}_1 - P_2 \subset \overline{P} - P_3$ telle que $C^k(\mathfrak{U}) - E^k(\mathfrak{U}) \to \Gamma^{k-1}(\mathfrak{U})$. Donc $D^k(\mathfrak{U}) - E^k(\mathfrak{U}) \to \Gamma^{k-1}(\mathfrak{U})$, d'où $\Gamma^{k-1} \sim 0$ dans $(\overline{P} - P_3)$, c.q.f.d.

18.2. Supposons la validité des axiomes G^k et H^{k-1}. Soit $a \in R - S$. On doit prouver que $\beta_k(a, R) = 0$. D'après l'axiome G^k, chaque entourage P suffisamment petit de a possède la propriété suivantes: chaque (k, R)-cycle absolu dans $\overline{P}$ est ~ 0 dans R. Il suffit de prouver que $\beta_k(a, P; R) = 0$ pour chaque tel entourage P. Soit P_1 un entourage de P tel que $\overline{P}_1 \subset P$. D'après l'axiome H^{k-1}, chaque entourage P_3 suffisamment petit de a possède la propriété suivante: Si Γ^{k-1} est un $(k-1, R)$-cycle absolu dans $(\overline{P}_1 - P_1)$ et si $\Gamma^{k-1} \sim 0$ dans $\overline{P}_1$, on a $\Gamma^{k-1} \sim 0$ dans $(\overline{P} - P_3)$. Il suffit de prouver que $\beta_k(P_3, P; R) = 0$ pour chaque tel entourage P_3. Soit C^k un (k, R)-cycle mod $(R - P)$. On doit prouver que $C^k \sim 0$ mod $(R - P_3)$.

Soit Φ la famille complète de réseaux qui se déduit de la famille N définie

[14] On doit remarquer que, si R est une $V_0^n(S)$, R vérifie toujours l'axiome H^{-1}, ce qui résulte p. ex. de *Variétés*, 12.2.

dans *Homologie*, IV, 2 en y remplaçant R_1, R_2, R_3, α, α_1, α_2, α_3, resp. par $\overline{P}_1$, $R - P_1$, $\overline{P}_1 - P_1$, $R - P$, 0, $R - P$, 0. Pour chaque $\mathfrak{U} \in \Phi$ on peut poser (v. *Homologie*, IV, 6, où on remplace n par $k - 1$) $C^k(\mathfrak{U}) = C_1^k(\mathfrak{U}) - C_2^k(\mathfrak{U})$, où $C_1^k(\mathfrak{U}) \subset \overline{P}_1$, $C_2^k(\mathfrak{U}) \subset R - P_1$. Posons $FC_1^k(\mathfrak{U}) = \Gamma^{k-1}(\mathfrak{U})$. Les $\Gamma^{k-1}(\mathfrak{U})$ définissent (*Homologie*, IV, 12) un $(k-1, R)$-cycle absolu Γ^{k-1} dans $(\overline{P}_1 - P_1)$ tel que $C_1^k(\mathfrak{U}) \to \Gamma^{k-1}(\mathfrak{U})$ pour chaque $\mathfrak{U} \in \Phi$, d'où $\Gamma^{k-1} \sim 0$ dans $\overline{P}_1$. On a donc $\Gamma^{k-1} \sim 0$ dans $(\overline{P} - P_3)$. Donc il existe pour chaque $\mathfrak{U} \in \Phi$ une $(k, \mathfrak{U})$-chaîne $D^k(\mathfrak{U})$ dans $(\overline{P} - P_3)$ telle que $D^k(\mathfrak{U}) \to \Gamma^{k-1}(\mathfrak{U})$, d'où $C_1^k(\mathfrak{U}) - D^k(\mathfrak{U}) \to 0$. Soit $\mathfrak{U} \in \Phi$; la famille Φ étant complète, il existe un affinement $\mathfrak{V} \in \Phi$ de $\mathfrak{U}$ normal rel. aux cycles absolus dans $\overline{P}$; soit $\pi = \mathrm{Pr}(\mathfrak{V}, \mathfrak{U})$. On a

$$C_1^k(\mathfrak{V}) \subset \overline{P}_1 \subset \overline{P},\ D^k(\mathfrak{V}) \subset \overline{P} - P_3 \subset \overline{P},\ C_1^k(\mathfrak{V}) - D^k(\mathfrak{V}) \to 0.$$

Donc $\pi C_1^k(\mathfrak{V}) - \pi D^k(\mathfrak{V})$ est un $(k, \mathfrak{U})$-cycle absolu essentiel dans $\overline{P}$, d'où $\pi C_1^k(\mathfrak{V}) - \pi D^k(\mathfrak{V}) \sim 0$. Donc il existe une $(k+1, \mathfrak{U})$-chaîne $M^{k+1}(\mathfrak{U})$ telle que $M^{k+1}(\mathfrak{U}) \to \pi C_1^k(\mathfrak{V}) - \pi D^k(\mathfrak{V})$. Comme $C^k(\mathfrak{V}) = C_1^k(\mathfrak{V}) - C_2^k(\mathfrak{V})$, on a

$$M^{k+1}(\mathfrak{V}) \to \pi C^k(\mathfrak{V}) + \pi C_2^k(\mathfrak{V}) - \pi D^k(\mathfrak{V}).$$

Or $C_2^k(\mathfrak{V}) \subset R - P_1 \subset R - P_3$, $D^k(\mathfrak{V}) \subset \overline{P} - P_3 \subset R - P_3$. Par suite $\pi C^k(\mathfrak{V}) \sim 0 \bmod (R - P_3)$. D'autre part $\pi C^k(\mathfrak{V}) \sim C^k(\mathfrak{U}) \bmod (R - P) \subset R - P_3$. Donc $C^k(\mathfrak{U}) \sim 0 \bmod (R - P_3)$ pour chaque $\mathfrak{U} \in \Phi$. La famille Φ étant complète, on a $C^k \sim 0 \bmod (R - P_3)$. c.q.f.d.

19. En tenant compte de 14, on déduit de 9, 13 et 18 d'après *Variétés*, 65 et 66 les théorèmes suivants.

19.1. *Soit* $0 \leqq p \leqq \dfrac{n-1}{2}$. *Une* $V_p^n(S)$ *vérifie les axiomes* I^k(n°15) *pour* $0 \leqq k \leqq p$.

19.2. *Soit* $0 \leqq p \leqq n/2 - 1$. *Une* $V_p^n(S)$ *vérifie les axiomes* II^k(n°15) *pour* $0 \leqq k \leqq p - 1$.

19.3. *Soit* $\dfrac{n-1}{2} \leqq p \leqq n$. *Une* $V_p^n(S)$ *est une* $V_n^n(S)$.

20. Soient M_1 et M_2 deux modules (*Homologie*, I, 1). D'après *M. Pontragin*,[15]) nous dirons que M_1 et M_2 sont *duels* (*primitifs*) si l'on a défini le produit $\alpha, \beta \in \mathfrak{R}$ pour $\alpha \in M_1$ et $\beta \in M_2$ jouissant des propriétés suivantes: 1° $(\alpha_1 + \alpha_2)\beta = \alpha_1\beta + \alpha_2\beta$ pour $\alpha_1 \in M_1$, $\alpha_2 \in M_1$, $\beta \in M_2$; 2° $\alpha(\beta_1 + \beta_2) = \alpha\beta_1 + \alpha\beta_2$ pour $\alpha \in M_1$, $\beta_1 \in M_2$, $\beta_2 \in M_2$; 3° $(r\alpha)\beta = \alpha(r\beta) = r(\alpha\beta)$ pour $r \in \mathfrak{R}$, $\alpha \in M_1$, $\beta \in M_2$; 4° si $\alpha \in M_1$ et si $\alpha\eta = 0$ pour chaque $\eta \in M_2$, on a $\alpha = 0$; 5° si $\beta \in M_2$ et si $\xi\beta = 0$ pour chaque $\xi \in M_1$, on a $\beta = 0$.

20.1. *Soient* M_1 *et* M_2 *deux modules duels. Soient* $\alpha_1, \alpha_2, \cdots, \alpha_m$ *des éléments de* M_1 *en nombre fini tels que* $\sum_{i=1}^{m} r_i \alpha_i = 0$ $(r_i \in \mathfrak{R})$ *entraîne que* $r_1 = \cdots = r_m = 0$.

---

[15]) Math. Annalen, t. 105, 1931, pp. 165–205. V. *Variétés*, 61.

Il existe des éléments $\beta_1, \beta_2, \cdots, \beta_m$ *de* M_2 *tels que* $\alpha_i\beta_j = \delta_{ij}$ pour $1 \leqq i \leqq m$, $1 \leqq j \leqq m$, δ_{ij} étant le symbole de Kronecker: $\delta_{ii} = 1$, $\delta_{ij} = 0$ pour $i \neq j$.

DÉMONSTRATION. I. Soit d'abord $m = 1$. On a $\alpha_1 \in M$, $\alpha_1 \neq 0$. D'après 20, 4° il existe un élément $\eta \in M_2$ tel que $\alpha_1\eta = r \neq 0$. Posons $\beta_1 = 1/r\ \eta$. Alors $\alpha_1\beta_1 = 1$ d'après 20,3°.

II. Soit $m \geqq 2$ et supposons que le théorème soit vrai pour $m - 1$. Donc il existe des éléments $\beta_1', \cdots, \beta_{m-1}'$ de M_2 tels que $\alpha_i\beta_j' = \delta_{ij}$ pour $1 \leqq i \leqq m - 1$, $1 \leqq j \leqq m - 1$. Soit $\alpha_m\beta_i' = r_i (1 \leqq i \leqq m - 1)$. Comme $\alpha_m - \sum_{i=1}^{m-1} r_i\alpha_i \neq 0$, d'après I il existe un élément β_m' de M_2 tel que $(\alpha_m - \sum_{i=1}^{m-1} r_i\alpha_i)\beta_m' = 1$. Posons $\alpha_i\beta_m' = s_i (1 \leqq i \leqq m - 1)$ Il suffit de poser

$$\beta_i = \beta_i' - r_i\beta_m' + r_i \sum_{j=1}^{m-1} s_j\beta_j' (1 \leqq i \leqq m - 1), \qquad \beta_m = \beta_m' - \sum_{j=1}^{m-1} s_j\beta_j'$$

pour que l'on ait $\alpha_i\beta_j = \delta_{ij}$ $(1 \leqq i \leqq m, 1 \leqq j \leqq m)$.

21. Rappelons encore le théorème de dualité sous la forme dont nous allons nous servir au Chap. V (*Variétés*, 64):

Soit $0 \leqq p \leqq n$. *Soit* $t = \min(p, n - p)$. *Soit* R *une* $V_t^n(S)$ *orientable* mod S. *Le* $p^{\text{ème}}$ *groupe de Betti* $\mathfrak{G}_p$ *de l'espace* R mod S *est duel au* $(n - p)^{\text{ème}}$ *groupe de Betti bicompact* $\mathfrak{H}_{n-p}$ *de l'espace* $R - S$.

Les éléments de $\mathfrak{G}_p$ sont les (p, R)-cycles mod S, deux tels cycles C^p et D^p étant considérés comme égaux si et seulement si $C^p \sim D^p$ mod S. Les éléments de $\mathfrak{H}_{n-p}$ sont les $(n - p, R)$-cycles absolus situés dans un sous-ensemble bicompact arbitraire de $R - S$, deux tels cycles Γ^{n-p} et Δ^{n-p} étant considérés comme égaux si et seulement s'il existe un sous-ensemble bicompact A de $R - S$ tel que $\Gamma^{n-p} \sim \Delta^{n-p}$ dans A.

IV.

22. Soit R un espace topologique. Soit B un sous-ensemble ouvert de R. Soit C^k un (k, R)-cycle absolu $(k = 0, 1, 2, \cdots)$. Nous disons que C^k *est situé à l'intérieur de* B s'il existe un ensemble $F \subset B$ fermé dans R et tel que $C^k \subset F$; nous disons que $C^k \sim 0$ *à l'intérieur de* B s'il existe un ensemble F fermé dans R et tel que $C^k \sim 0$ dans F.

23. Soit R un espace topologique. Soit A un sous-ensemble fermé de R. Soit a un point donné de A. Soit $k = 0, 1, 2, \cdots$. Les lettres U, V, W désignent des entourages de a dans R.

Soit $U \supset V$. Soit $\mathfrak{A}_k(V, U; R - A)$ l'ensemble de tous les (k, R)-cycles absolus Γ^k situés à l'intérieur de $V - A$; pour $k = 0$ il faut supposer encore que $J(\Gamma^k) = 0$. Deux éléments Γ^k et Δ^k de $\mathfrak{A}_k(V, U; R - A)$ seront considérés comme égaux si et seulement si $\Gamma^k \sim \Delta^k$ à l'intérieur de $U - A$. L'ensemble

dans *Homologie*, IV, 2 en y remplaçant R_1, R_2, R_3, α, α_1, α_2, α_3, resp. par $\overline{P}_1$, $R - P_1$, $\overline{P}_1 - P_1$, $R - P$, 0, $R - P$, 0. Pour chaque $\mathfrak{U} \in \Phi$ on peut poser (v. *Homologie*, IV, 6, où on remplace n par $k - 1$) $C^k(\mathfrak{U}) = C_1^k(\mathfrak{U}) - C_2^k(\mathfrak{U})$, où $C_1^k(\mathfrak{U}) \subset \overline{P}_1$, $C_2^k(\mathfrak{U}) \subset R - P_1$. Posons $FC_1^k(\mathfrak{U}) = \Gamma^{k-1}(\mathfrak{U})$. Les $\Gamma^{k-1}(\mathfrak{U})$ définissent (*Homologie*, IV, 12) un $(k-1, R)$-cycle absolu Γ^{k-1} dans $(\overline{P}_1 - P_1)$ tel que $C_1^k(\mathfrak{U}) \to \Gamma^{k-1}(\mathfrak{U})$ pour chaque $\mathfrak{U} \in \Phi$, d'où $\Gamma^{k-1} \sim 0$ dans $\overline{P}_1$. On a donc $\Gamma^{k-1} \sim 0$ dans $(\overline{P} - P_3)$. Donc il existe pour chaque $\mathfrak{U} \in \Phi$ une $(k, \mathfrak{U})$-chaîne $D^k(\mathfrak{U})$ dans $(\overline{P} - P_3)$ telle que $D^k(\mathfrak{U}) \to \Gamma^{k-1}(\mathfrak{U})$, d'où $C_1^k(\mathfrak{U}) - D^k(\mathfrak{U}) \to 0$. Soit $\mathfrak{U} \in \Phi$; la famille Φ étant complète, il existe un affinement $\mathfrak{V} \in \Phi$ de $\mathfrak{U}$ normal rel. aux cycles absolus dans $\overline{P}$; soit $\pi = \mathrm{Pr}(\mathfrak{V}, \mathfrak{U})$. On a

$$C_1^k(\mathfrak{V}) \subset \overline{P}_1 \subset \overline{P},\ D^k(\mathfrak{V}) \subset \overline{P} - P_3 \subset \overline{P},\ C_1^k(\mathfrak{V}) - D^k(\mathfrak{V}) \to 0.$$

Donc $\pi C_1^k(\mathfrak{V}) - \pi D^k(\mathfrak{V})$ est un $(k, \mathfrak{U})$-cycle absolu essentiel dans $\overline{P}$, d'où $\pi C_1^k(\mathfrak{V}) - \pi D^k(\mathfrak{V}) \sim 0$. Donc il existe une $(k+1, \mathfrak{U})$-chaîne $M^{k+1}(\mathfrak{U})$ telle que $M^{k+1}(\mathfrak{U}) \to \pi C_1^k(\mathfrak{V}) - \pi D^k(\mathfrak{V})$. Comme $C^k(\mathfrak{V}) = C_1^k(\mathfrak{V}) - C_2^k(\mathfrak{V})$, on a

$$M^{k+1}(\mathfrak{V}) \to \pi C^k(\mathfrak{V}) + \pi C_2^k(\mathfrak{V}) - \pi D^k(\mathfrak{V}).$$

Or $C_2^k(\mathfrak{V}) \subset R - P_1 \subset R - P_3$, $D^k(\mathfrak{V}) \subset \overline{P} - P_3 \subset R - P_3$. Par suite $\pi C^k(\mathfrak{V}) \sim 0 \bmod (R - P_3)$. D'autre part $\pi C^k(\mathfrak{V}) \sim C^k(\mathfrak{U}) \bmod (R - P) \subset R - P_3$. Donc $C^k(\mathfrak{U}) \sim 0 \bmod (R - P_3)$ pour chaque $\mathfrak{U} \in \Phi$. La famille Φ étant complète, on a $C^k \sim 0 \bmod (R - P_3)$. c.q.f.d.

19. En tenant compte de 14, on déduit de 9, 13 et 18 d'après *Variétés*, 65 et 66 les théorèmes suivants.

19.1. *Soit* $0 \leqq p \leqq \dfrac{n-1}{2}$. *Une* $V_p^n(S)$ *vérifie les axiomes* I^k(n°15) *pour* $0 \leqq k \leqq p$.

19.2. *Soit* $0 \leqq p \leqq n/2 - 1$. *Une* $V_p^n(S)$ *vérifie les axiomes* II^k(n°15) *pour* $0 \leqq k \leqq p - 1$.

19.3. *Soit* $\dfrac{n-1}{2} \leqq p \leqq n$. *Une* $V_p^n(S)$ *est une* $V_n^n(S)$.

20. Soient M_1 et M_2 deux modules (*Homologie*, I, 1). D'après *M. Pontragin*,[15]) nous dirons que M_1 et M_2 sont *duels* (*primitifs*) si l'on a défini le produit $\alpha, \beta \in \mathfrak{R}$ pour $\alpha \in M_1$ et $\beta \in M_2$ jouissant des propriétés suivantes: 1° $(\alpha_1 + \alpha_2)\beta = \alpha_1\beta + \alpha_2\beta$ pour $\alpha_1 \in M_1$, $\alpha_2 \in M_1$, $\beta \in M_2$; 2° $\alpha(\beta_1 + \beta_2) = \alpha\beta_1 + \alpha\beta_2$ pour $\alpha \in M_1$, $\beta_1 \in M_2$, $\beta_2 \in M_2$; 3° $(r\alpha)\beta = \alpha(r\beta) = r(\alpha\beta)$ pour $r \in \mathfrak{R}$, $\alpha \in M_1$, $\beta \in M_2$; 4° si $\alpha \in M_1$ et si $\alpha\eta = 0$ pour chaque $\eta \in M_2$, on a $\alpha = 0$; 5° si $\beta \in M_2$ et si $\xi\beta = 0$ pour chaque $\xi \in M_1$, on a $\beta = 0$.

20.1. *Soient* M_1 *et* M_2 *deux modules duels. Soient* $\alpha_1, \alpha_2, \cdots, \alpha_m$ *des éléments de* M_1 *en nombre fini tels que* $\sum_{i=1}^{m} r_i \alpha_i = 0$ $(r_i \in \mathfrak{R})$ *entraîne que* $r_1 = \cdots = r_m = 0$.

---

[15]) Math. Annalen, t. 105, 1931, pp. 165–205. V. *Variétés*, 61.

Il existe des éléments $\beta_1, \beta_2, \cdots, \beta_m$ *de* M_2 *tels que* $\alpha_i\beta_j = \delta_{ij}$ pour $1 \leqq i \leqq m$, $1 \leqq j \leqq m$, δ_{ij} étant le symbole de Kronecker: $\delta_{ii} = 1$, $\delta_{ij} = 0$ pour $i \neq j$.

DÉMONSTRATION. I. Soit d'abord $m = 1$. On a $\alpha_1 \in M$, $\alpha_1 \neq 0$. D'après 20, 4° il existe un élément $\eta \in M_2$ tel que $\alpha_1\eta = r \neq 0$. Posons $\beta_1 = 1/r\ \eta$. Alors $\alpha_1\beta_1 = 1$ d'après 20,3°.

II. Soit $m \geqq 2$ et supposons que le théorème soit vrai pour $m - 1$. Donc il existe des éléments $\beta_1', \cdots, \beta_{m-1}'$ de M_2 tels que $\alpha_i\beta_j = \delta_{ij}$ pour $1 \leqq i \leqq m-1$, $1 \leqq j \leqq m-1$. Soit $\alpha_m\beta_i' = r_i (1 \leqq i \leqq m-1)$. Comme $\alpha_m - \sum_{i=1}^{m-1} r_i\alpha_i \neq 0$, d'après I il existe un élément β_m' de M_2 tel que $(\alpha_m - \sum_{i=1}^{m-1} r_i\alpha_i)\beta_m' = 1$. Posons $\alpha_i\beta_m' = s_i (1 \leqq i \leqq m-1)$ Il suffit de poser

$$\beta_i = \beta_i' - r_i\beta_m' + r_i \sum_{j=1}^{m-1} s_j\beta_j' (1 \leqq i \leqq m-1), \qquad \beta_m = \beta_m' - \sum_{j=1}^{m-1} s_j\beta_j'$$

pour que l'on ait $\alpha_i\beta_j = \delta_{ij}$ $(1 \leqq i \leqq m, 1 \leqq j \leqq m)$.

21. Rappelons encore le théorème de dualité sous la forme dont nous allons nous servir au Chap. V (*Variétés*, 64):

Soit $0 \leqq p \leqq n$. *Soit* $t = \min(p, n-p)$. *Soit* R *une* $V_i^n(S)$ *orientable* mod S. *Le* $p^{\text{ème}}$ *groupe de Betti* $\mathfrak{G}_p$ *de l'espace* R mod S *est duel au* $(n-p)^{\text{ème}}$ *groupe de Betti bicompact* $\mathfrak{H}_{n-p}$ *de l'espace* $R - S$.

Les éléments de $\mathfrak{G}_p$ sont les (p, R)-cycles mod S, deux tels cycles C^p et D^p étant considérés comme égaux si et seulement si $C^p \sim D^p$ mod S. Les éléments de $\mathfrak{H}_{n-p}$ sont les $(n-p, R)$-cycles absolus situés dans un sous-ensemble bicompact arbitraire de $R - S$, deux tels cycles Γ^{n-p} et Δ^{n-p} étant considérés comme égaux si et seulement s'il existe un sous-ensemble bicompact A de $R - S$ tel que $\Gamma^{n-p} \sim \Delta^{n-p}$ dans A.

IV.

22. Soit R un espace topologique. Soit B un sous-ensemble ouvert de R. Soit C^k un (k, R)-cycle absolu $(k = 0, 1, 2, \cdots)$. Nous disons que C^k *est situé à l'intérieur de* B s'il existe un ensemble $F \subset B$ fermé dans R et tel que $C^k \subset F$; nous disons que $C^k \sim 0$ *à l'intérieur de* B s'il existe un ensemble F fermé dans R et tel que $C^k \sim 0$ dans F.

23. Soit R un espace topologique. Soit A un sous-ensemble fermé de R. Soit a un point donné de A. Soit $k = 0, 1, 2, \cdots$. Les lettres U, V, W désignent des entourages de a dans R.

Soit $U \supset V$. Soit $\mathfrak{A}_k(V, U; R - A)$ l'ensemble de tous les (k, R)-cycles absolus Γ^k situés à l'intérieur de $V - A$; pour $k = 0$ il faut supposer encore que $J(\Gamma^k) = 0$. Deux éléments Γ^k et Δ^k de $\mathfrak{A}_k(V, U; R - A)$ seront considérés comme égaux si et seulement si $\Gamma^k \sim \Delta^k$ à l'intérieur de $U - A$. L'ensemble

$\mathfrak{A}_k(V, U; R - A)$ est un module. Désignons par $\alpha_k(V, U; R - A)$ le rang de ce module si ce rang est fini; dans le cas contraire posons $\alpha_k(V, U; R - A) = \infty$.

Pour $W \subset V \subset U$ on a évidemment $\alpha_k(W, U; R - A) \leqq \alpha_k(V, U; R - A)$. Il en résulte que, l'entourage U de a étant donné, le nombre $\alpha_k(V, U; R - A)$ a une valeur fixe (indépendante de V) pour tous les voisinages $V \subset U$ de a suffisamment petits; désignons par $\alpha_k(a, U; R - A)$ cette valeur fixe.

Pour $W \subset V \subset U$ on a évidemment $\alpha_k(W, V; R - A) \geqq \alpha_k(W, U; R - A)$. On en déduit sans peine que $\alpha_k(a, V; R - A) \geqq \alpha_k(a, U; R - A)$ pour $V \subset U$, Par suite trois cas sont à distinguer:

1°. Il existe un nombre fini m $(= 0, 1, 2, \cdots)$ tel que $\alpha_k(a, U; R - A) = m$ pour tous les entourages U de a suffisamment petits; dans ce cas on pose $\alpha_k(a, R - A) = m$.

2°. Le nombre $\alpha_k(a, U; R - A)$ est fini pour les entourages U de a, mais si m est un nombre fini arbitrairement donné, on a $\alpha_k(a, U; R - A) > m$ pour tous les entourages U de a suffisamment petits; dans ce cas on pose $\alpha_k(a, R - A) = \omega$.

3°. On a $\alpha_k(a, U; R - A) = \infty$ pour tous les entourages U de a suffisamment petits; dans ce cas on pose $\alpha_k(a, R - A) = \infty$.

24. En général, le nombre $\alpha_k(a, R - A)$ n'est pas complètement déterminé par l'espace $(R - A) + (a)$, il faut connaître aussi l'espace R. Or on voit sans peine (v. *Homologie*, III, 3–11) que, l'espace $(R - A) + (a)$ et son point a étant donnés, pour connaître le nombre $\alpha_k(a, R - A)$ il suffit de savoir encore de chaque sous-ensemble de $R - A$ s'il est ou non fermé dans R. En particulier si l'espace R est bicompact, le nombre $\alpha_k(a, R - A)$ ne dépend que de l'espace $(R - A) + (a)$ (localement) et de son point a; car un sous-ensemble de $R - A$ est alors fermé dans R si et seulement s'il est bicompact. Il suffit même que l'espace R soit localement bicompact au point a, c'est-a-dire qu'il existe un entourage U de a tel que l'ensemble soit bicompact.

25. Les énoncés suivants sont faciles à vérifier:[16]

L'équation $\alpha_0(a, R - A) = m$ $(= 0, 1, 2, \cdots)$ *signifie que, si* U *est un entourage de* a *suffisamment petit, le point* a *appartient à la frontière de* $(m + 1)$ *constituants*[17] *de* $U - A$, *tandis que* a *n'appartient pas à la frontière de la somme de tous les autres constituants de* $U - A$. *Si l'espace* R *est régulier et localement connexe, l'équation* $\alpha_0(a, R - A) = m$ *signifie qu'il existe des entourages* U *de* a *arbitrairement petits tels que l'ensemble* $U - A$ *ait* $m + 1$ *composantes tandis que, pour chaque entourage* U *de* a *suffisamment petit, l'ensemble* $U - A$ *a plus que* m *composantes.*

L'équation $\alpha_0(a, R - A) = \omega$ *signifie*: 1° *si* U *est un entourage de* a *suffisamment*

---

[16] Le cas banal où A contient tout un entourage de a dans R y est tacitement exclu.

[17] Un *constituant* d'un sous-ensemble M de R est un sous-ensemble de M maximé (saturé) rel. à la propriété que chaque couple de ses points se laisse unir par un sous-ensemble de M connexe et fermé dans R.

petit, le point a appartient à la frontière d'un nombre fini de constituants de $U - A$, tandis que a n'appartient pas à la frontière de la somme de tous les autres constituants de a, 2° m étant un nombre fini arbitrairement donné, le point a appartient à la frontière de plus que m constituants de l'ensemble $U - A$ pour chaque entourage de a suffisamment petit. Si l'espace R est régulier et localement connexe, l'équation $\alpha_0(a, R - A) = \omega$ signifie: 1° il existe des entourages U de a arbitrairement petits tels que l'ensemble $U - A$ ait un nombre fini de composantes; 2° m étant un nombre fini arbitrairement donné, l'ensemble $U - A$ a plus que m composantes pour chaque entourage U de a suffisamment petit.

V.

26. *Soit R une $V_0^n(S)$. Soit A un sous-ensemble fermé de R. Soit $a \in A - S$. Alors $\alpha_0(a, R - A) \leqq \beta_{n-1}(a, A)$.*

DÉMONSTRATION. Le théorème est banal si $a \in A - \overline{R - A}$, car alors $\alpha_0(a, R - A) = 0$. Supposons donc que $a \in \overline{R - A}$. On voit sans peine qu'il suffit de démontrer l'énoncé suivant: Soit $m = 1, 2, 3, \cdots$; soit U un entourage de a suffisamment petit; soit $\alpha_0(a, U; R - A) \geqq m$; alors $\beta_{n-1}(a, U; A) \geqq m$.

Soit $U \subset R - S$ un entourage de a si petit que (v. 14) R soit orientable mod $(R - U)$. Soit $\alpha_0(a, U; R - A) \geqq m$. On doit prouver que $\beta_{n-1}(V, U; A) \geqq m$ pour chaque entourage $V \subset U$ de a.

D'après 7, 6° on a $\beta_n(a, R) = 1$, d'où $\beta_n(a, V; R) \leqq 1$. Donc il existe un entourage $W \subset V$ de a tel que $\beta_n(W, V; R) \leqq 1$. Comme $a \in \overline{R - A}$, on a $W - A \neq 0$.

Comme $\alpha_0(a, U; R - A) \geqq m$, on a $\alpha_0(W, U; R - A) \geqq m$. Donc il existe des $(0, R)$-cycles absolus $\Gamma_i^0 (1 \leqq i \leqq m)$ situés à l'intérieur de $W - A$, tels que $J(\Gamma_i^0) = 0$ et tels que l'homologie $\sum_{i=1}^{m} r_i \Gamma_i^0 \sim 0$ à l'intérieur de $U - A (r_i \in \Re)$ entraîne que $r_1 = \cdots = r_m = 0$. D'après 21 (v. aussi 20.1), où on pose $p = n$ et où on remplace S par $A + (R - U)$ (v. 8 et 12), il existe des (n, R)-cycles $E_i^n (1 \leqq i \leqq m)$ mod $A + (R - U)$ tels que $E_i^n \Gamma_j^0 = \delta_{ij}$ pour $1 \leqq i \leqq m$, $1 \leqq j \leqq m$. Soit G^n un (n, R)-cycle principal mod $(R - U)$.

Supposons que $rG^n + \sum_{i=1}^{m} r_i E_i^n \sim 0 \bmod A + (R - W) (r \in \Re, r_i \in \Re)$. Il en résulte (*Variétés*, 56, 4°) que $(rG^n + \sum_{i=1}^{m} r_i E_i^n)$. $\Gamma_j^0 = 0$ pour $1 \leqq j \leqq m$. Or on a $G^n \Gamma_j^0 = 0$, car $J(\Gamma_j^0) = 0$. (V. *Variétés*, 19.1, 4° et 57.) Donc $\sum_{i=1}^{m} r_i E_i^n \Gamma_j^0 = 0$ pour $1 \leqq j \leqq m$, d'où $r_1 = \cdots = r_m = 0$, donc $rG^n \sim 0 \bmod A + (R - W)$. Comme $W - A \neq 0$, le cycle G^n n'est pas $\sim 0 \bmod A + (R - W)$ (v. 11). Donc $r = 0$.

Nous venons de prouver que les $G^n, E_1^n, \cdots, E_m^n$ ne sont liés par aucune homologie mod $A + (R - W)$. Donc il existe un réseau $\mathfrak{U}_1$ tel que l'homologie

$rG^n(\mathfrak{U}_1) + \sum_{i=1}^{m} r_i E_i^n(\mathfrak{U}_1) \sim 0 \bmod A + (R - W)(r \in \mathfrak{R}, r_i \in \mathfrak{R})$ entraîne que $r = r_1 = \cdots = r_m = 0$; d'après 11 on peut supposer que $G^n(\mathfrak{U}_1)$ n'est pas ~ 0 mod $(R - W)$. Soit $\mathfrak{U}_2$ un affinement de $\mathfrak{U}_1$ normal rel. aux cycles mod $(R - V)$. Soit $\mathfrak{U}_3$ un affinement de $\mathfrak{U}_2$ normal rel. aux cycles mod $(A - U)$ dans A. Déterminons un affinement $\mathfrak{U}_4$ de $\mathfrak{U}_3$ d'après 4, en y posant $\varphi = A$, $\chi = R - U$. Soit $\pi_{21} = \text{Pr.}(\mathfrak{U}_2, \mathfrak{U}_1), \cdots, \pi_{31} = \pi_{21} \cdot \pi_{32}, \cdots$.

Les E_i^n étant des (n, R)-cycles mod $A + (R - U)$, il existe des $(n - 1, \mathfrak{U}_4)$-chaînes $C_i^{n-1}(\mathfrak{U}_4) \subset A$ et $D_i^{n-1}(\mathfrak{U}_4) \subset R - U$ telles que $E_i^n(\mathfrak{U}_4) \to C_i^{n-1}(\mathfrak{U}_4) + D_i^{n-1}(\mathfrak{U}_4)$. Donc $FC_i^{n-1}(\mathfrak{U}_4) = -FD_i^{n-1}(\mathfrak{U}_4) \subset R - U$. D'autre part $FC_i^{n-1}(\mathfrak{U}_4) \subset A$. Donc $FC_i^{n-1}(\mathfrak{U}_4) \subset A - U$ d'après la définition de $\mathfrak{U}_4$. En vertu de la définition de $\mathfrak{U}_3$, les $\pi_{42} C_i^{n-1}(\mathfrak{U}_4) (1 \leqq i \leqq m)$ sont donc des $(n - 1, \mathfrak{U}_2)$-cycles essentiels mod $(A - U)$ dans A. Reste à prouver que les $\pi_{42} C_i^{n-1}(\mathfrak{U}_4)$ ne sont liés par aucune homologie mod $(A - V)$ dans A, car alors (v. *Homologie*, II, 28) on a $\beta_{n-1}(V, U; A) \geqq m$, c.q.f.d.

Soit donc $\pi_{42} \sum_{i=1}^{m} r_i C_i^{n-1}(\mathfrak{U}_4) \sim 0 \bmod (A - V)$ dans $A (r_i \in \mathfrak{R})$; on doit prouver que $r_1 = \cdots = r_m = 0$. L'homologie qui vient d'être écrite signifie qu'il existe une $(n, \mathfrak{U}_2)$-chaîne $N^n(\mathfrak{U}_2) \subset A$ et une $(n - 1, \mathfrak{U}_2)$-chaîne $T^{n-1}(\mathfrak{U}_2) \subset A - V$ telles que $N^n(\mathfrak{U}_2) \to \pi_{42} \sum_{i=1}^{m} r_i C_i^{n-1}(\mathfrak{U}_4) + T^{n-1}(\mathfrak{U}_2)$. Comme $E_i^n(\mathfrak{U}_4) \to C_i^{n-1}(\mathfrak{U}_4) + D_i^{n-1}(\mathfrak{U}_4)$, on a

$$\pi_{42} \sum_{i=1}^{m} r_i E_i^n(\mathfrak{U}_4) - N^n(\mathfrak{U}_2) \to \pi_{42} \sum_{i=1}^{m} r_i D_i^{n-1}(\mathfrak{U}_4) - T^{n-1}(\mathfrak{U}_2) \subset R - V.$$

En vertu de la définition de $\mathfrak{U}_2$, on en déduit que $\pi_{41} \sum_{i=1}^{m} r_i E_i^n(\mathfrak{U}_4) \to \pi_{21} N^n(\mathfrak{U}_2)$ est un $(n, \mathfrak{U}_1)$-cycle essentiel mod $(R - V)$. Or nous savons d'une part que $G^n(\mathfrak{U}_1)$ n'est pas ~ 0 mod $(R - W)$, d'autre part que $\beta_n(W, V; R) \leqq 1$. On en déduit qu'il existe un nombre $r \in \mathfrak{R}$ tel que

$$rG^n(\mathfrak{U}_1) + \pi_{41} \sum_{i=1}^{m} r_i E_i^n(\mathfrak{U}_4) - \pi_{21} N^n(\mathfrak{U}_2) \sim 0 \bmod (R - W),$$

ce qui entraîne que

$$rG^n(\mathfrak{U}_1) + \pi_{41} \sum_{i=1}^{m} r_i E_i^n(\mathfrak{U}_4) \sim 0 \bmod A + (R - W).$$

Or $\pi_{41} E_i^n(\mathfrak{U}_4) \sim E_i^n(\mathfrak{U}_1) \bmod A + (R - V) \subset A + (R - W)$. Donc $rG^n(\mathfrak{U}_1) + \sum_{i=1}^{m} r_i E_i^n(\mathfrak{U}_1) \sim 0 \bmod A + (R - W)$. D'après la définition de $\mathfrak{U}_1$, il en résulte que $r_1 = \cdots = r_m = 0$, c.q.f.d.

27. *Soit* $0 \leqq p \leqq n-2$. *Soit* $q = n-p-1$. *Soit* $t = \min(p+1, q)$. *Soit* R *une* $V_i^n(S)$. *Soit* A *un sous-ensemble fermé de* R. *Soit* $a \in R - S$. *Soit* $\beta_{p+1}(a, R) = 0$.[18] *Alors* $\alpha_q(a, R-A) \leqq \beta_p(a, A)$.

DÉMONSTRATION. On voit sans peine qu'il suffit de démontrer l'énoncé suivant: Soit $m = 1, 2, \cdots$; soit U un entourage de a suffisamment petit, soit $\alpha_q(a, U; R-A) \geqq m$; alors $\beta_p(a, U; R-A) \geqq m$.

Soit $U \subset R - S$ un entourage de a si petit que (v. 14) R soit orientable mod $(R-U)$. Soit $\alpha_q(a, U; R-A) \geqq m$. On doit prouver que $\beta_p(V, U; A) \geqq m$ pour chaque entourage $V \subset U$ de a.

Puisque $\beta_{p+1}(a, R) = 0$, on a $\beta_{p+1}(a, V; R) = 0$. Donc il existe un entourage $W \subset V$ de a tel que $\beta_{p+1}(W, V; R) = 0$. Comme $\alpha_q(a, U; R-A) \geqq m$, on a $\alpha_q(W, U; R-A) \geqq m$. Donc il existe des (q, R)-cycles absolus $\Gamma_i^q (1 \leqq i \leqq m)$ situés à l'intérieur de $W - A$ et tels que l'homologie $\sum_{i=1}^{m} r_i \Gamma_i^q \sim 0$ à l'intérieur de A $(r_i \in \mathfrak{R})$ entraîne que $r_1 = \cdots = r_m = 0$.

D'après 21 (v. aussi 20.1), où on remplace p par $p+1$ et S par $A + (R-U)$ (v. 12 et 16) il existe des $(p+1, R)$-cycles $E_i^{p+1} (1 \leqq i \leqq m)$ mod $A + (R-U)$ tels que $E_i^{p+1} \Gamma_i^q = \delta_{ij}$ $(1 \leqq i \leqq m, 1 \leqq j \leqq m)$.

Soit $\sum_{i=1}^{m} r_i E_i^{p+1} \sim 0 \bmod A + (R-W)$ $(r_i \in \mathfrak{R})$. Alors $\sum_{i=1}^{m} r_i E_i^{p+1} \Gamma_j^q \sim 0$ (*Variétés*, 56, 4°), d'où $r_1 = r_2 = \cdots = r_m = 0$. Donc les $E_1^{p+1}, \cdots, E_m^{p+1}$ ne sont liés par aucune homologie mod $A + (R-W)$. Donc il existe un réseau $\mathfrak{U}_1$ tel que l'homologie $\sum_{i=1}^{m} r_i E_i^{p+1}(\mathfrak{U}_1) \sim 0 \bmod A + (R-W)$ entraîne que $r_1 = \cdots = r_m = 0$. Soit $\mathfrak{U}_2$ un affinement de $\mathfrak{U}_1$ normal rel. aux cycles mod $(R-V)$.

Soit $\mathfrak{U}_3$ un affinement de $\mathfrak{U}_2$ normal rel. aux cycles mod $(A-U)$ dans A. Déterminons un affinement $\mathfrak{U}_4$ de $\mathfrak{U}_3$ d'après 4, en y posant $\varphi = A$, $\chi = R - U$. Soit $\pi_{21} = \text{Pr.}(\mathfrak{U}_2, \mathfrak{U}_1), \cdots, \pi_{31} = \pi_{21} \cdot \pi_{32}, \cdots$.

Les E_i^{p+1} étant des $(p+1, R)$-cycles mod $A + (R-U)$, il existe des $(p, \mathfrak{U}_4)$-chaînes $C_i^p(\mathfrak{U}_4) \subset A$ et $D_i^p(\mathfrak{U}_4) \subset R - U$ telles que $E_i^{p+1}(\mathfrak{U}_4) \to C_i^p(\mathfrak{U}_4) + D_i^p(\mathfrak{U}_4)$. Donc $FC_i^p(\mathfrak{U}_4) - FD_i^p(\mathfrak{U}_4) \subset R - U$. D'autre part $FC_i^p(\mathfrak{U}_4) \subset A$. Donc $FC_i^p(\mathfrak{U}_4) \subset A - U$ d'après la définition de $\mathfrak{U}_4$. En vertu de la définition de $\mathfrak{U}_3$, les $\pi_{42} C_i^p(\mathfrak{U}_4)$ $(1 \leqq i \leqq m)$ sont donc des $(p, \mathfrak{U}_2)$-cycles essentiels mod $(A-U)$ dans A. Reste à prouver que les $\pi_{42} C_i^p(\mathfrak{U}_4)$ ne sont liés par aucune homologie mod $(A-V)$ dans A, car alors (v. *Homologie*, II, 28) on a $\beta_p(V, U; A) \geqq m$, c.q.f.d. Soit donc $\pi_{42} \sum_{i=1}^{m} r_i C_i^p(\mathfrak{U}_4) \sim 0 \bmod (A-V)$ dans $A (r_i \in \mathfrak{R})$; on doit prouver que $r_1 = \cdots = r_m = 0$. L'homologie qui vient

[18] *Si* $p \leqq \dfrac{n-1}{2}$ cette hypothèse est une conséquence des—précédentes en vertu de 19.2.

d'être écrite signifie qu'il existe une $(p+1, \mathfrak{U}_2)$-chaîne $N^{p+1}(\mathfrak{U}_2) \subset A$ et une $(p, \mathfrak{U}_2)$-chaîne $T^p(\mathfrak{U}_2) \subset A - V$ telles que $N^{p+1}(\mathfrak{U}_2) \to \pi_{42} \sum_{i=1}^{m} r_i C_i^p(\mathfrak{U}_4) - T^p(\mathfrak{U}_2)$. Comme $E_i^{p+1}(\mathfrak{U}_4) \to C_i^p(\mathfrak{U}_4) + D_i^p(\mathfrak{U}_4)$, on a

$$\pi_{42} \sum_{i=1}^{m} r_i E_i^{p+1}(\mathfrak{U}_4) - N^{p+1}(\mathfrak{U}_2) \to \pi_{42} \sum_{i=1}^{m} r_i D_i^p(\mathfrak{U}_4) - T^p(\mathfrak{U}_2) \subset R - V.$$

En vertu de la définition de $\mathfrak{U}_2$, on en déduit que $\pi_{41} \sum_{i=1}^{m} r_i E_i^{p+1}(\mathfrak{U}_4) \to \pi_{21} N^{p+1}(\mathfrak{U}_2)$ est un $(p+1, \mathfrak{U}_4)$-cycle essentiel mod $(R - V)$. Or on a $\beta_{p+1}(W, V; R) = 0$. Donc $\pi_{41} \sum_{i=1}^{m} r_i E_i^{p+1}(\mathfrak{U}_4) - \pi_{21} N^{p+1}(\mathfrak{U}_2) \sim 0 \bmod (R - W)$, d'où $\pi_{41} \sum_{i=1}^{m} r_i E_i^{p+1}(\mathfrak{U}_4) \sim 0 \bmod A + (R - W)$. Or $\pi_{41} E_i^{p+1}(\mathfrak{U}_4) \sim E_i^{p+1}(\mathfrak{U}_1) \bmod A + (R - V) \subset A + (R - W)$. Donc $\sum_{i=1}^{m} r_i E_i^{p+1}(\mathfrak{U}_1) \sim 0 \bmod A + (R - W)$. D'après la définition de $\mathfrak{U}_1$, il en résulte que $r_1 = \cdots = r_m = 0$, c.q.f.d.

28. *Soit R une $V_0^q(S)$. Soit A un sous-ensemble fermé de R. Soit $a \in A - S$. Soit $\beta_{n-1}(a, R) = 0$. Alors $\alpha_0(a, R - A) \geqq \beta_{n-1}(a, A)$.*

DÉMONSTRATION. Le théorème est banal si $a \in A - \overline{R - A}$, car alors $\beta_{n-1}(a, A) = \beta_{n-1}(a, R) = 0$. Supposons donc que $a \in \overline{R - A}$. On voit sans peine qu'il suffit de démontrer l'énoncé suivant: soit $m = 1, 2, 3, \cdots$; soit U un entourage de a suffisamment petit; soit $\beta_{n-1}(a, U; A) \geqq m$; alors il existe un entourage $V \subset U$ de a tel que $\alpha_0(a, V; R - A) \geqq m$.

Soit $U \subset R - S$ un entourage de a si petit que (v. 14) R soit orientable mod $(R - U)$. Soit G^n un (n, R)-cycle principal mod $(R - U)$. Soit $\beta_{n-1}(a, U; A) \geqq m$. Comme $\beta_{n-1}(a, R) = 0$, on a $\beta_{n-1}(a, U; R) = 0$; donc il existe un entourage $V \subset U$ de a tel que $\beta_{n-1}(V, U; R) = 0$. On doit prouver que $\alpha_0(W, V; R - A) \geqq m$ pour chaque entourage $W \subset V$ de a. Comme $a \in \overline{R - A}$, on a $W - A \neq 0$.

Puisque $\beta_{n-1}(a, U; A) \geqq m$, on a $\beta_{n-1}(W, U; A) \geqq m$. Donc il existe des $(n-1, R)$-cycles $C_i^{n-1} (1 \leqq i \leqq m) \bmod (A - U)$ dans A tels que l'homologie $\sum_{i=1}^{m} r_i C_i^{n-1} \sim 0 \bmod (A - W)$ dans A $(r_i \in \mathfrak{R})$ entraîne que $r_1 = \cdots = r_m = 0$.

Soit $\mathfrak{U}_1$ un réseau tel que l'homologie $\sum_{i=1}^{m} r_i C_i^{n-1}(\mathfrak{U}_1) \sim 0 \bmod (A - W)$ dans A entraîne que $r_1 = \cdots = r_m = 0$. Déterminons un affinement $\mathfrak{U}_2$ de $\mathfrak{U}_1$ d'après 4, en y posant $\varphi = A$, $\chi = R - W$. Puisque $W - A \neq 0$, on peut

supposer (v. 11) que $G^n(\mathfrak{U}_2)$ n'est pas $\sim 0 \bmod A + (R - W)$. Soit $\mathfrak{U}_3$ un affinement de $\mathfrak{U}_2$ normal rel. aux cycles mod $A + (R - V)$. Soit $\pi_{21} = \text{Pr.}(\mathfrak{U}_2, \mathfrak{U}_1)$, $\pi_{32} = \text{Pr.}(\mathfrak{U}_3, \mathfrak{U}_2)$, $\pi_{31} = \pi_{21} \cdot \pi_{32}$. Les C_i^{n-1} étant des $(n-1, R)$-cycles mod $(R - U)$, on a $C_i^{n-1} \sim 0 \bmod (R - V)$, car $\beta_{n-1}(V, U; R) = 0$. Donc il existe des $(n, \mathfrak{U}_3)$-chaînes $E_i^n(\mathfrak{U}_3)$ et des $(n-1, \mathfrak{U}_3)$-chaînes $D_i^{n-1}(\mathfrak{U}_3) \subset R - V$ telles que $E_i^n(\mathfrak{U}_3) \to C_i^{n-1}(\mathfrak{U}_3) - D_i^{n-1}(\mathfrak{U}_3)$, D'après la définition de $\mathfrak{U}_3$, il en résulte que les $\pi_{32} E_i^n(\mathfrak{U}_3)$ sont des $(n, \mathfrak{U}_2)$-cycles essentiels mod $A + (R - V)$.

Si l'on a $rG^n(\mathfrak{U}_2) + \pi_{32} \sum_{i=1}^{m} r_i E_i^n(\mathfrak{U}_3) \sim 0 \bmod A + (R - W) (r \in \mathfrak{R}, r_i \in \mathfrak{R})$, il existe des $\mathfrak{U}_2$-chaînes $M^{n+1}(\mathfrak{U}_2)$, $N_1^n(\mathfrak{U}_2) \subset A$ et $N_2^n(\mathfrak{U}_2) \subset R - W$ telles que

$$M^{n+1}(\mathfrak{U}_2) \to rG^n(\mathfrak{U}_2) + \pi_{32} \sum_{i=1}^{m} r_i E_i^n(\mathfrak{U}_3) - N_1^n(\mathfrak{U}_2) - N_2^n(\mathfrak{U}) \to 0.$$

Comme $E_i^n(\mathfrak{U}_3) \to C_i^{n-1}(\mathfrak{U}_3) - D_i^{n-1}(\mathfrak{U}_3)$, il en résulte que

$$\pi_{32} \sum_{i=1}^{m} r_i C_i^{n-1}(\mathfrak{U}_3) = \pi_{32} \sum_{i=1}^{m} r_i D_i^{n-1}(\mathfrak{U}_3) - rFG^n(\mathfrak{U}_2) + FN_1^n(\mathfrak{U}_2) + FN_2^n(\mathfrak{U}_2),$$

d'où on déduit que la chaîne

(1) $FN_1^n(\mathfrak{U}_2) - \pi_{32} \sum_{i=1}^{m} r_i C_i^{n-1}(\mathfrak{U}_3)$ est située dans $(R - W)$. La chaîne (1) étant aussi située dans A, il résulte de la définition de $\mathfrak{U}_2$ qu'elle est située dans $(A - W)$. Donc $\pi_{32} \sum_{i=1}^{m} r_i C_i^{n-1}(\mathfrak{U}_3) \sim 0 \bmod (A - W)$ dans A, d'où $\pi_{31} \sum_{i=1}^{m} r_i C_i^{n-1}(\mathfrak{U}_3) \sim 0 \bmod (A - W)$ dans A. Or $\pi_{31} C_i^{n-1}(\mathfrak{U}_3) \sim C_i^{n-1}(\mathfrak{U}_1) \bmod (A - U)$ dans A. Comme $U \supset W$, il en résulte que $\sum_{i=1}^{m} r_i C_i^{n-1}(\mathfrak{U}_1) \sim 0 \bmod (A - W)$ dans A, d'où $r_1 = \cdots = r_m = 0$, d'après la définition de $\mathfrak{U}_1$. Donc $rC^n(\mathfrak{U}_2) \sim 0 \bmod A + (R - W)$, ce qui entraîne que $r = 0$.

Nous venons de prouver que les

(2) $G^n(\mathfrak{U}_2), \pi_{32} E_1^n(\mathfrak{U}_3), \cdots, \pi_{32} E_m^n(\mathfrak{U}_3)$ ne sont liés par aucune homologie mod $A + (R - W)$. Or les (2) sont des $(n, \mathfrak{U}_2)$-cycles essentiels mod $A + (R - V)$. Soient

(3) $G^n, E_1^n, \cdots, E_m^n$ des (n, R)-cycles mod $A + (R - V)$ attachés aux (2) selons *Homologie*, II, 28. Evidemment les (3) ne sont liés par aucune homologie mod $A + (R - W)$. D'après 21 (v. aussi 20.1), où on pose $p = n$ et où on remplace S par $A + (R - W)$ (v. 8 et 12), il existe des $(0, R)$-cycles $\Gamma_i^\circ (0 \leqq i \leqq m)$ situés à l'intérieur de $W - A$ et tels que

$$G^n \Gamma_0^\circ = 1, G^n \Gamma_i^\circ = 0 \ (1 \leqq i \leqq m), E_i^n \Gamma_j^\circ = \delta_{ij} \ (1 \leqq i \leqq m, 1 \leqq j \leqq m).$$

Comme $G^n \Gamma_i^\circ = 0$ $(1 \leqq i \leqq m)$, on voit sans peine (v. *Variétés*, 19.1, 4° et 57) que $J(\Gamma_i^\circ) = 0$ pour $1 \leqq i \leqq m$. Soit $\sum_{i=1}^{m} r_i \Gamma_i^\circ \sim 0$ à l'intérieur de $V - A$ $(r_i \epsilon \Re)$. Comme E_j^n $(1 \leqq j \leqq m)$ est un (n, R)-cycle mod $A + (R - V)$, il résulte de *Variétés*, 56, 4° que $E_j^n \sum_{i=1}^{m} r_i \Gamma_i^\circ = 0$, d'où $r_1 = \cdots = r_m = 0$. Donc les Γ_i^0 $(1 \leqq i \leqq m)$ sont des $(0, R)$-cycles absolus située à l'intérieur de $W - A$, ils ne sont liés par aucune homologie à l'intérieur de $W - A$, et on a $J(\Gamma_i^\circ) = 0$ $(1 \leqq i \leqq m)$. Donc $\alpha_0(W, V; R - A) \geqq m$, c.q.f.d.

29. *Soit* $0 \leqq p \leqq n - 2$. *Soit* $q = n - p - 1$. *Soit* $t = \min(p + 1, q)$. *Soit* R *une* $V_t^n(S)$. *Soit* A *un sous-ensemble fermé de* R. *Soit* $a \epsilon A - S$. *Soit* $\beta_p(a, R) = 0$.[19] *Alors* $\alpha_q(a, R - A) \geqq \beta_p(a, A)$.

DÉMONSTRATION. On voit sans peine qu'il suffit de démontrer l'énoncé suivant: Soit $m = 1, 2, 3, \cdots$; soit U un entourage de a suffisamment petit; soit $\beta_p(a, U; A) \geqq m$; alors il existe un entourage $V \subset U$ de a tel que $\alpha_q(a, V; R - A) \geqq m$.

Soit $U \subset R - S$ un entourage de a si petit que (v. 14) R soit orientable mod $(R - U)$. Soit $\beta_p(a, U; A) \geqq m$. Comme $\beta_p(a, R) = 0$, on a $\beta_p(a, U; R) = 0$; donc il existe un entourage $V \subset U$ de a tel que $\beta_p(V, U; R) = 0$. On doit prouver que $\alpha_q(W, V; R - A) \geqq m$ pour chaque entourage $W \subset V$ de a.

Puisque $\beta_p(a, U; A) \geqq m$ on a $\beta_p(W, U; A) \geqq m$. Donc il existe des (p, R)-cycles $C_i^p (1 \leqq i \leqq m)$ mod $(A - U)$ dans A tels que l'homologie $\sum_{i=1}^{m} r_i C_i^p \sim 0 \bmod (A - W)$ dans A $(r_i \epsilon \Re)$ entraîne que $r_1 = \cdots = r_m = 0$. Soit $\mathfrak{U}_1$ un réseau tel que l'homologie $\sum_{i=1}^{m} r_i C_i^p(\mathfrak{U}_1) \sim 0 \bmod (A - W)$ dans A entraîne que $r_1 = \cdots = r_m = 0$. Déterminons un affinement $\mathfrak{U}_2$ de $\mathfrak{U}_1$ d'après 4, en y posant $\varphi = A$, $\chi = R - W$. Soit $\mathfrak{U}_3$ un affinement de $\mathfrak{U}_2$ normal rel. aux cycles mod $A + (R - V)$. Soit $\pi_{21} = \text{Pr.}(\mathfrak{U}_2, \mathfrak{U}_1)$, $\pi_{32} = \text{Pr.}(\mathfrak{U}_3, \mathfrak{U}_2)$, $\pi_{31} = \pi_{21}\pi_{32}$.

Les C_i^p étant des (p, R)-cycles mod $(R - U)$, on a $C_i^p \sim 0$ mod $(R - V)$, car $\beta_p(V, U; R) = 0$. Donc il existe des $(p + 1, \mathfrak{U}_3)$-chaînes $E_i^{p+1}(\mathfrak{U}_3)$ et des $(p, \mathfrak{U}_3)$-chaînes $D_i^p(\mathfrak{U}_3) \subset R - V$ telles que $E_i^{p+1}(\mathfrak{U}_3) \to C_i^p(\mathfrak{U}_3) - D_i^p(\mathfrak{U}_3)$. D'après la définition de $\mathfrak{U}_3$, il en résulte que les $\pi_{32} E_i^{p+1}(\mathfrak{U}_3)$ sont des $(p + 1, \mathfrak{U}_2)$-cycles essentiels mod $A + (R - V)$.

Si l'on a $\pi_{32} \sum_{i=1}^{m} r_i E_i^{p+1}(\mathfrak{U}_3) \sim 0$ mod $A + (R - W)(r_i \epsilon \Re)$, il existe des $\mathfrak{U}_2$-chaînes $M^{p+2}(\mathfrak{U}_2)$, $N_1^{p+1}(\mathfrak{U}_2) \subset A$ et $N_2^{p+1}(\mathfrak{U}_2) \subset R - W$ telles que

$$M^{p+2}(\mathfrak{U}_2) \to \pi_{32} \sum_{i=1}^{m} r_i E_i^{p+1}(\mathfrak{U}_0) - N_1^{p+1}(\mathfrak{U}_2) - N_2^{p+1}(\mathfrak{U}_2) \to 0.$$

[19] Si $p \leqq n/2$, cette hypothèse est une conséquence des précedentes en vertu de 19.2.

Comme $E_i^{p+1}(\mathfrak{U}_3) \to C_i^p(\mathfrak{U}_3) - D_i^p(\mathfrak{U}_3)$, il en résulte que

$$\pi_{32} \sum_{i=1}^{m} r_i C_i^p(\mathfrak{U}_3) = \pi \sum_{i=1}^{m} r_i D_i^p(\mathfrak{U}_3) + FN_1^{p+1}(\mathfrak{U}_2) + FN_2^{p+1}(\mathfrak{U}_2),$$

d'où on déduit que la chaine

(1) $FN_1^{p+1}(\mathfrak{U}_2) - \pi_{32} \sum_{i=1}^{m} r_i C_i^p(\mathfrak{U}_3)$ est située dans $(R - W)$. La chaîne (1) étant aussi située dans A, il résulte de la définition de $\mathfrak{U}_2$ qu'elle est située dans $(A - W)$. Donc $\pi_{32} \sum_{i=1}^{m} r_i C_i^p(\mathfrak{U}_3) \sim 0 \bmod (A - W)$ dans A, d'où $\pi_{31} \sum_{i=1}^{m} r_i C_i^p(\mathfrak{U}_3) \sim 0 \bmod (A - W)$ dans A. Or $\pi_{31} C_i^p(\mathfrak{U}_3) \sim C_i^p(\mathfrak{U}_1) \bmod (A - U)$ dans A. Comme $U \supset W$, il en résulte que $\sum_{i=1}^{m} r_i C_i^p(\mathfrak{U}_1) \sim 0 \bmod (A - W)$ dans A, d'où $r_1 = \cdots = r_m = 0$ d'après la définition de U_1.

Nous venons de prouver que les

(2) $\pi_{32} E_1^{p+1}(\mathfrak{U}_3) \cdots , \pi_{32} E_m^{p+1}(\mathfrak{U}_3)$ ne sont liés par aucune homologie mod $A + (R - W)$. Or les (2) sont des $(p + 1, \mathfrak{U}_2)$-cycles essentiels mod $A + (R - V)$. Soient

(3) $E_1^{p+1}, \cdots , E_m^{p+1}$ des $(p + 1, R)$-cycles mod $A + (R - V)$ attachés aux (2) selon *Homologie*, II, 28. Evidemment les (3) ne sont liés par aucune homologie mod $A + (R - W)$. D'après 21 (v. aussi 20.1), où on remplace p par $p + 1$ et S par $A + (R - W)$ (v. 12 et 16), il existe des (q, R)-cycles absolus $\Gamma_i^q (1 \leqq i \leqq m)$ situés à l'intérieur de $(W - A)$ et tels que $E_i^{p+1} \Gamma_j^q = \delta_{ij}$ $(1 \leqq i \leqq m, 1 \leqq j \leqq m)$. Soit $\sum_{i=1}^{m} r_i \Gamma_i^q \sim 0$ à l'intérieur de $V - A (r_i \epsilon \Re)$. Comme $E_j^{p+1} (1 \leqq j \leqq m)$ est un $(p + 1, R)$-cycle mod $A + (R - V)$, il résulte de *Variétés*, 56, 4° que $E_j^{p+1} \sum_{i=1}^{m} r_i \Gamma_i^q = 0$, d'où $r_1 = \cdots = r_m = 0$. Donc les $\Gamma_i^q (1 \leqq i \leqq m)$ sont des (q, R) absolus situés à l'intérieur de $W - A$ et ne liés par aucune homologie à l'intérieur de $V - A$. Donc $\alpha_q(W, V; R - A) \geqq m$, c.q.f.d.

BRNO, CZECHOSLOVAKIA.

ANNALS OF MATHEMATICS
Vol. 35, No. 3, July, 1934

SOME GROUP-THEORETIC CONSIDERATIONS IN ALGEBRAIC GEOMETRY

BY J. A. TODD

(Received May 7, 1934)

In a series of recent papers[1] Severi has developed at great length the notion of *equivalence* between sets of points on an algebraic surface, and defined certain aggregates of these sets which he calls *series of equivalence*. His definition[2] of the latter by intersections of the surface with loci belonging to a continuous family is framed in such a way that any continuous aggregate of sets belonging to such a series is itself a series of equivalence. Thus the notion of equivalence, rather than that of series of equivalence is the fundamental one. One aim of this note is to show that this concept of equivalence arises in a very simple way from *group-theoretic* considerations of a general character, and that when regarded from this angle the whole theory becomes very simple. The other purpose of the note is to emphasize these group notions which though implicit in many of the concepts of algebraic geometry, (particularly these concerning *virtual* entities) do not appear to have been explicitly formulated hitherto. When viewed from our standpoint such ideas as linear systems and continuous systems appear as particular applications of a general principle.

I should like to express my thanks to Professor Lefschetz for much helpful advice and criticism in the preparation of this note.

Consider an algebraic V_d, and the aggregate of all algebraic V_k, reducible and irreducible, which lie upon it. If A and B are two such V_k then their sum C is also a member of the aggregate, and may be expressed in the form $C = A + B$. This defines an operation of addition among the V_k, and this operation clearly has the property that $A + B = A + B_1$ only when B and B_1 coincide. If $C = A + B$ we may say that A is the *difference* between C and B, and write $A = C - B$; such a relation being interpreted as meaning the same as $C = A + B$. The V_k are the elements of an *abelian semi-group* $\mathfrak{S}$ with addition as the rule of combination.

We propose to extend $\mathfrak{S}$ into an *abelian group*. To do this we introduce, corresponding to each element A of $\mathfrak{S}$, an element A^* with the property that

$$A + A^* = A^* + A = 0.$$

We adjoin these new elements to $\mathfrak{S}$ and assume that they obey the usual rules of addition. We have, clearly, $(A + B)^* = A^* + B^*$. If, now, $A + B = C$,

[1] (A). Comm. Mat. Helvetici, 4 (1932), 268. (B). Mem. R. Acc. Italia, 3 (1932), 5. (C). Ibid., 4 (1933), 71. (D). Rend. Acc. Lincei (6), 17 (1933), 419, 491, 597.

[2] (D). 422.

then by adding B^* to each side we get $A = C + B^*$. Hence if we write $-B$ instead of B^* the use of the negative sign agrees with the previous use in those cases where a meaning was attached to it. Consequently, by adjunction of these new elements $\mathfrak{S}$ has been converted into an abelian group $\mathfrak{A}$, in which A^* (or $-A$) is the inverse of A. Any element of $\mathfrak{A}$ can be expressed (in many ways) in the form $C - B$, where C and B belong to $\mathfrak{S}$. If C does not contain B as a part the element is called virtual. *The elements of $\mathfrak{A}$ are all the V_k, effective and virtual, on V_d.*

Consider now any subgroup $\mathfrak{H}$ of $\mathfrak{A}$. This defines a division of the elements of $\mathfrak{A}$ into cosets, and these cosets themselves form an abelian group $\mathfrak{G}$ simply isomorphic with the factor group $\mathfrak{A}/\mathfrak{H}$. The elements of $\mathfrak{G}$ are aggregates of V_k (virtual or effective) such that the difference between any two V_k of the aggregate belongs to $\mathfrak{H}$. The subgroup $\mathfrak{H}$ defines a relation of *equivalence* between the V_k which belong to any one of the cosets, this relation, in view of the group properties, being symmetrical, reflexive, and transitive, and being preserved by addition. The elements of $\mathfrak{H}$ are those elements which are equivalent to zero in this sense. This notion is clearly invariant for birational transformations of V_d.

The theory of linear systems of V_{d-1} on V_d, of continuous systems of curves on a surface, or of series of equivalent sets of points on a surface, are all included in the foregoing theory, for a suitable choice of $\mathfrak{H}$. We shall illustrate this, very briefly in the first two cases, and in more detail in the last case, since this method of approach seems much simpler than Severi's. (For example, the elaborate theory of series of equivalence on reducible curves[3] can be dispensed with entirely.)

Suppose, then, that $k = d - 1$. Let $\mathfrak{H}$ consist of these elements of the form $A - B$, where A, B are such that, for some C, $A + C$ and $B + C$ form the complete intersection of V_d with two hypersurfaces of a linear pencil. It is very easily seen that these elements actually form a group. The elements of $\mathfrak{G}$ in this case are the *complete linear systems* (effective and virtual) of V_{d-1} on V_d, the notion of equivalence defined by $\mathfrak{H}$ is that of linear equivalence.

Again, let $d = 2$, $k = 1$. Let $\mathfrak{H}$ consist of all V_1 on V_2 which are mapped on the Riemannian of the surface by cycles which are homologous to zero. The elements of $\mathfrak{G}$ are then *complete continuous systems* of curves on the surface, and the relation of equivalence is that of algebraic equivalence in the sense of Severi, or of topological equivalence in the sense of Lefschetz.[4]

Let $d = 2$, $k = 0$, so that we are dealing with sets of points on an algebraic surface. We shall show how Severi's notion of equivalence can be defined very easily in terms of the group concepts discussed above. We define a series of intersection on the surface F as the series of sets of points of intersection of two curves which move in linear systems on F, together with possible limit points

[3] (B). 6–27.

[4] L'Analysis Situs et la géométrie algébrique, (Paris, 1924), 80.

arising when the two curves have a common part. We now define the group $\mathfrak{H}$ as follows: it contains all sets of points of the form $A - B$, where A, B are elements of $\mathfrak{S}$ such that, for some C, $A + C$ and $B + C$ belong to a series of intersection; and all finite sums of such sets; i.e. the general element of $\mathfrak{H}$ has the form $(A_1 - B_1) + (A_2 - B_2) + \cdots + (A_i - B_i)$. These sets evidently form a group. The elements of $\mathfrak{G}$ are then aggregates of sets mutually equivalent in the sense of Severi, where we have included in $\mathfrak{G}$ certain virtual sets of points as well as the effective ones which Severi considers. To prove this it is sufficient to show that two effective sets A, B which belong to the same element of $\mathfrak{G}$ are equivalent in the sense of Severi, and conversely. For brevity, write $A \equiv B$ if A, B are equivalent in the sense of Severi, and $A \sim B$ if A, B belong to the same element of $\mathfrak{G}$. Suppose first that $A \sim B$. Then $A - B$ is an element of $\mathfrak{H}$ and we have

$$A - B = (A_1 - B_1) + \cdots + (A_i - B_i), \tag{1}$$

where for each $j = 1, 2, \cdots, i$ there exists a C_j such that $A_j + C_j$, $B_j + C_j$ belong to a series of intersection. Hence (as follows easily from the definition of series of equivalence) $A_j + C_j \equiv B_j + C_j$. Hence $\Sigma A_j + \Sigma C_j \equiv \Sigma B_j + \Sigma C_j$. So $\Sigma A_j \equiv \Sigma B_j$. But, from (1), $A + \Sigma B_j = B + \Sigma A_j$. Hence, as all the sets concerned are effective, $A \equiv B$.[5] Suppose, conversely, that $A \equiv B$. Then[6] if K denotes the set of points common to A, B, the sets $A - K$, $B - K$ are either characteristic sets of a net of curves on F, or become such on adding sets C, D which are equivalent on some curve Γ on F. This means that $C = C_1 + \cdots + C_h$, $C = D_1 + \cdots + D_h$, where C_j, D_j belong to the same irreducible part of Γ and form sets of the same linear series on this curve. We have thus the relations

$$A - K + C \sim B - K + D,$$

while

$$C_j \sim D_j \qquad (j = 1, \cdots, h).$$

Hence from the group properties, $A \sim B$ and our theorem is established.

Severi's principal theorems[7] concerning equivalent sets state essentially that they possess the group property, and from our point of view these results are immediate consequences of the definition.

The theory may be extended inductively to the more general case of V_k on V_d. The definition of $\mathfrak{H}$ will be similar to the previous one, where by a series of intersection we shall mean an aggregate of V_k in which the V_{d-1} of a linear system meet a set of equivalent V_{k+1}. This leads to an inductive definition of equivalence between V_k, since for V_{d-1} the corresponding notion is that of linear equivalence. It is then a simple matter to prove that the variable intersections of a family of equivalent V_k with a family of equivalent $V_{k'}$, form a family of equivalent $V_{k+k'-d}$.

PRINCETON UNIVERSITY.

[5] (D). 491.
[6] (D). 420.
[7] (D). 491–494; (C). 87.